PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C12N 15/54, 15/29, 15/82, 9/12, A01H 5/00 | A2

(11) International Publication Number: **WO 00/70059**

(43) International Publication Date: 23 November 2000 (23.11.00)

(21) International Application Number: PCT/US00/11687

(22) International Filing Date: 28 April 2000 (28.04.00)

(30) Priority Data:

60/134,292	14 May 1999 (14.05.99)	US
60/142,996	8 July 1999 (08.07.99)	US

(71) Applicant *(for all designated States except US)*: PIONEER HI-BRED INTERNATIONAL, INC. [US/US]; 800 Capital Square, 400 Locust Street, Des Moines, IA 50309 (US).

(72) Inventor; and

(75) Inventor/Applicant *(for US only)*: HELENTJARIS, Timothy, G. [US/US]; 2960 NW 73rd Lane, Ankeny, IA 50021 (US).

(74) Agents: VARLEY, Karen, K. et al.; 7100 N.W. 62nd Avenue, Darwin Building, Johnston, IA 50131-1000 (US).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

Without international search report and to be republished upon receipt of that report.

(54) Title: SIGNAL TRANSDUCTION GENES AND METHODS OF USE

(57) Abstract

The invention provides isolated signal transduction nucleic acids and their encoded proteins. The present invention provides methods and compositions relating to altering signal transduction gene expression levels in plants. The invention further provides recombinant expression cassettes, host cells, transgenic plants, and antibody compositions.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

SIGNAL TRANSDUCTION GENES AND METHODS OF USE

TECHNICAL FIELD

The present invention relates generally to plant molecular biology. More specifically, it relates to nucleic acids and methods for modulating their expression in plants.

BACKGROUND OF THE INVENTION

Cells are constantly bombarded by external signals that regulate their growth, differentiation, and stress level. To respond properly to these signals, eukaryotic cells assemble cascades of highly conserved protein kinases (mitogen-activated protein kinases, MAPKs, and their activator kinases), which form the central elements of signal transduction pathways that lead to and activate transcription factors in the nucleus and other effectors throughout the cell. (Elion, E.A. *Routing MAP Kinase Cascades. Science*, Vol. 281 (11 September 1998), pp. 1625-1626)

Plants' sessile nature requires strategies to adapt to changes in their local environment. Responses to signals such as pathogen attack, herbivore feeding, or changes in light, temperature, or nutrient availability often involve phytohormones such as ethylene. MAP kinase modules appear to be involved in ethylene signaling and auxin-induced cell proliferation. Upstream and downstream regulatory components suggest that MAP kinases may serve as a central point in a signal transduction network where different upstream pathways converge and may distribute the signals to different downstream targets. (Jonak, C., et al. *MAP kinases: universal multi-purpose signaling tools. Plant Molecular Biology* Vol. 24 (1994), pp. 407-416.)

DESCRIPTION OF THE INVENTION

Overview

Generally, it is the object of the present invention to provide nucleic acids and proteins relating to signal transduction genes. It is an object of the present invention to provide transgenic plants comprising the nucleic acids of the present invention, and methods for modulating, in a transgenic plant, expression of the nucleic acids of the present invention.

Therefore, in one aspect the present invention relates to an isolated nucleic acid comprising a member selected from the group consisting of (a) a polynucleotide having a

specified sequence identity to a polynucleotide encoding a polypeptide of the present invention; (b) a polynucleotide which is complementary to the polynucleotide of (a); and, (c) a polynucleotide comprising a specified number of contiguous nucleotides from a polynucleotide of (a) or (b). The isolated nucleic acid can be DNA.

In other aspects the present invention relates to: 1) recombinant expression cassettes, comprising a nucleic acid of the present invention operably linked to a promoter, 2) a host cell into which has been introduced the recombinant expression cassette, and 3) a transgenic plant comprising the recombinant expression cassette. The host cell and plant are optionally from maize, wheat, rice, or soybean.

Unless otherwise stated, the polynucleotide and polypeptide sequences identified in the accompanying Sequence Listing represent polynucleotides and polypeptides of the present invention. A nucleic acid of the present invention comprises a polynucleotide of the present invention. A protein of the present invention comprises a polypeptide of the present invention.

The present invention provides utility in such exemplary applications as regulating expression of phytohormones, including ethylene, auxins, cytokinins, and gibberellin, in order to effect developmental changes in plants and provide control of plant response to environmental stresses. Various preferred embodiments are disclosed throughout the specification.

Definitions

Units, prefixes, and symbols may be denoted in their SI accepted form. Unless otherwise indicated, nucleic acids are written left to right in 5' to 3' orientation; amino acid sequences are written left to right in amino to carboxy orientation, respectively. Numeric ranges recited within the specification are inclusive of the numbers defining the range and include each integer within the defined range. Amino acids may be referred to herein by either their commonly known three letter symbols or by the one-letter symbols recommended by the IUPAC-IUBMB Nomenclature Commission. Nucleotides, likewise, may be referred to by their commonly accepted single-letter codes. Unless otherwise provided for, software, electrical, and electronics terms as used herein are as defined in The New IEEE Standard Dictionary of Electrical and Electronics Terms (5th edition, 1993). The terms defined below are more fully defined by reference to the specification as a whole. Section headings provided throughout the specification are not limitations to the various objects and embodiments of the present invention.

By "amplified" is meant the construction of multiple copies of a nucleic acid sequence or multiple copies complementary to the nucleic acid sequence using at least one of the nucleic acid sequences as a template. Amplification systems include the polymerase chain reaction (PCR) system, ligase chain reaction (LCR) system, nucleic acid sequence based amplification (NASBA, Cangene, Mississauga, Ontario), Q-*Beta* Replicase systems, transcription-based amplification system (TAS), and strand displacement amplification (SDA). See, e.g., *Diagnostic Molecular Microbiology: Principles and Applications*, D. H. Persing *et al.*, Ed., American Society for Microbiology, Washington, D.C. (1993). The product of amplification is termed an amplicon.

As used herein, "antisense orientation" includes reference to a duplex polynucleotide sequence that is operably linked to a promoter in an orientation where the antisense strand is transcribed. The antisense strand is sufficiently complementary to an endogenous transcription product such that translation of the endogenous transcription product is often inhibited.

By "encoding" or "encoded", with respect to a specified nucleic acid, is meant comprising the information for translation into the specified protein. A nucleic acid encoding a protein may comprise non-translated sequences (e.g., introns) within translated regions of the nucleic acid, or may lack such intervening non-translated sequences (e.g., as in cDNA). The information by which a protein is encoded is specified by the use of codons. Typically, the amino acid sequence is encoded by the nucleic acid using the "universal" genetic code. However, variants of the universal code, such as are present in some plant, animal, and fungal mitochondria, the bacterium *Mycoplasma capricolum*, or the ciliate *Macronucleus*, may be used when the nucleic acid is expressed therein.

When the nucleic acid is prepared or altered synthetically, advantage can be taken of known codon preferences of the intended host where the nucleic acid is to be expressed. For example, although nucleic acid sequences of the present invention may be expressed in both monocotyledonous and dicotyledonous plant species, sequences can be modified to account for the specific codon preferences and GC content preferences of monocotyledons or dicotyledons as these preferences have been shown to differ (Murray *et al. Nucl. Acids Res.* 17: 477-498 (1989)). Thus, the maize preferred codon for a particular amino acid may be derived from known gene sequences from maize. Maize codon usage for 28 genes from maize plants is listed in Table 4 of Murray *et al.*, *supra*.

As used herein "full-length sequence" in reference to a specified polynucleotide or its encoded protein means having the entire amino acid sequence of, a native (non-

synthetic), endogenous, biologically (e.g., structurally or catalytically) active form of the specified protein. Methods to determine whether a sequence is full-length are well known in the art including such exemplary techniques as northern or western blots, primer extension, S1 protection, and ribonuclease protection. See, e.g., *Plant Molecular Biology: A Laboratory Manual*, Clark, Ed., Springer-Verlag, Berlin (1997). Comparison to known full-length homologous (orthologous and/or paralogous) sequences can also be used to identify full-length sequences of the present invention. Additionally, consensus sequences typically present at the 5' and 3' untranslated regions of mRNA aid in the identification of a polynucleotide as full-length. For example, the consensus sequence ANNNNAUGG, where the underlined codon represents the N-terminal methionine, aids in determining whether the polynucleotide has a complete 5' end. Consensus sequences at the 3' end, such as polyadenylation sequences, aid in determining whether the polynucleotide has a complete 3' end.

As used herein, "heterologous" in reference to a nucleic acid is a nucleic acid that originates from a foreign species, or, if from the same species, is substantially modified from its native form in composition and/or genomic locus by human intervention. For example, a promoter operably linked to a heterologous structural gene is from a species different from that from which the structural gene was derived, or, if from the same species, one or both are substantially modified from their original form. A heterologous protein may originate from a foreign species or, if from the same species, is substantially modified from its original form by human intervention.

By "host cell" is meant a cell which contains a vector and supports the replication and/or expression of the vector. Host cells may be prokaryotic cells such as *E. coli*, or eukaryotic cells such as yeast, insect, amphibian, or mammalian cells. Preferably, host cells are monocotyledonous or dicotyledonous plant cells. A particularly preferred monocotyledonous host cell is a maize host cell.

The term "introduced" includes reference to the incorporation of a nucleic acid into a eukaryotic or prokaryotic cell where the nucleic acid may be incorporated into the genome of the cell (e.g., chromosome, plasmid, plastid or mitochondrial DNA), converted into an autonomous replicon, or transiently expressed (e.g., transfected mRNA). The term includes such nucleic acid introduction means as "transfection", "transformation" and "transduction".

The term "isolated" refers to material, such as a nucleic acid or a protein, which is: (1) substantially or essentially free from components which normally accompany or

interact with it as found in its natural environment. The isolated material optionally comprises material not found with the material in its natural environment; or (2) if the material is in its natural environment, the material has been synthetically altered or synthetically produced by deliberate human intervention and/or placed at a different location within the cell. The synthetic alteration or creation of the material can be performed on the material within or apart from its natural state. For example, a naturally-occurring nucleic acid becomes an isolated nucleic acid if it is altered or produced by non-natural, synthetic methods, or if it is transcribed from DNA which has been altered or produced by non-natural, synthetic methods. The isolated nucleic acid may also be produced by the synthetic re-arrangement ("shuffling") of a part or parts of one or more allelic forms of the gene of interest. Likewise, a naturally-occurring nucleic acid (e.g., a promoter) becomes isolated if it is introduced to a different locus of the genome. Nucleic acids which are "isolated," as defined herein, are also referred to as "heterologous" nucleic acids. See, e.g., Compounds and Methods for Site Directed Mutagenesis in Eukaryotic Cells, Kmiec, U.S. Patent No. 5,565,350; *In Vivo* Homologous Sequence Targeting in Eukaryotic Cells; Zarling *et al.*, WO 93/22443; Methods for In Vitro Recombination, Stemmer, U.S. Patent No. 5,605,793.

As used herein, "nucleic acid" includes reference to a deoxyribonucleotide or ribonucleotide polymer, or chimeras thereof, in either single- or double-stranded form, and unless otherwise limited, encompasses known analogues having the essential nature of natural nucleotides in that they hybridize to single-stranded nucleic acids in a manner similar to naturally occurring nucleotides (e.g., peptide nucleic acids).

Unless otherwise stated, a "signal transduction nucleic acid" is a nucleic acid of the present invention and means a nucleic acid comprising a polynucleotide of the present invention (a "signal transduction polynucleotide") encoding a signal transduction polypeptide. A "signal transduction gene" is a gene of the present invention and refers to a full-length signal transduction polynucleotide.

By "nucleic acid library" is meant a collection of isolated DNA or RNA molecules which comprise and substantially represent the entire transcribed fraction of a genome of a specified organism, tissue, or of a cell type from that organism. Construction of exemplary nucleic acid libraries, such as genomic and cDNA libraries, is taught in standard molecular biology references such as Berger and Kimmel, *Guide to Molecular Cloning Techniques, Methods in Enzymology,* Vol. 152, Academic Press, Inc., San Diego, CA (Berger); Sambrook *et al., Molecular Cloning - A Laboratory Manual,* 2nd ed., Vol. 1-3 (1989); and

Current Protocols in Molecular Biology, F.M. Ausubel *et al.*, Eds., Current Protocols, a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc. (1994).

As used herein "operably linked" includes reference to a functional linkage between a promoter and a second sequence, wherein the promoter sequence initiates and mediates transcription of the DNA sequence corresponding to the second sequence. Generally, operably linked means that the nucleic acid sequences being linked are contiguous and, where necessary to join two protein coding regions, contiguous and in the same reading frame.

As used herein, the term "plant" includes reference to whole plants, plant parts or organs (e.g., leaves, stems, roots, etc.), plant cells, seeds and progeny of same. Plant cell, as used herein, further includes, without limitation, cells obtained from or found in: seeds, suspension cultures, embryos, meristematic regions, callus tissue, leaves, roots, shoots, gametophytes, sporophytes, pollen, and microspores. Plant cells can also be understood to include modified cells, such as protoplasts, obtained from the aforementioned tissues. The class of plants which can be used in the methods of the invention is generally as broad as the class of higher plants amenable to transformation techniques, including both monocotyledonous and dicotyledonous plants. A particularly preferred plant is *Zea mays*.

As used herein, "polynucleotide" includes reference to a deoxyribopolynucleotide, ribopolynucleotide, or chimeras or analogs thereof that have the essential nature of a natural deoxy- or ribo- nucleotide in that they hybridize, under stringent hybridization conditions, to substantially the same nucleotide sequence as naturally occurring nucleotides and/or allow translation into the same amino acid(s) as the naturally occurring nucleotide(s). A polynucleotide can be full-length or a subsequence of a native or heterologous structural or regulatory gene. Unless otherwise indicated, the term includes reference to the specified sequence as well as the complementary sequence thereof. Thus, DNAs or RNAs with backbones modified for stability or for other reasons are "polynucleotides" as that term is intended herein. Moreover, DNAs or RNAs comprising unusual bases, such as inosine, or modified bases, such as tritylated bases, to name just two examples, are polynucleotides as the term is used herein. It will be appreciated that a great variety of modifications have been made to DNA and RNA that serve many useful purposes known to those of skill in the art. The term polynucleotide as it is employed herein embraces such chemically, enzymatically or metabolically modified forms of polynucleotides, as well

as the chemical forms of DNA and RNA characteristic of viruses and cells, including among other things, simple and complex cells.

The terms "polypeptide", "peptide" and "protein" are used interchangeably herein to refer to a polymer of amino acid residues. The terms apply to amino acid polymers in which one or more amino acid residue is an artificial chemical analogue of a corresponding naturally occurring amino acid, as well as to naturally occurring amino acid polymers. The essential nature of such analogues of naturally occurring amino acids is that, when incorporated into a protein, that protein is specifically reactive to antibodies elicited to the same protein but consisting entirely of naturally occurring amino acids. The terms "polypeptide", "peptide" and "protein" are also inclusive of modifications including, but not limited to, glycosylation, lipid attachment, sulfation, gamma-carboxylation of glutamic acid residues, hydroxylation and ADP-ribosylation. Further, this invention contemplates the use of both the methionine-containing and the methionine-less amino terminal variants of the protein of the invention.

As used herein "promoter" includes reference to a region of DNA upstream from the start of transcription and involved in recognition and binding of RNA polymerase and other proteins to initiate transcription. A "plant promoter" is a promoter capable of initiating transcription in plant cells whether or not its origin is a plant cell. Exemplary plant promoters include, but are not limited to, those that are obtained from plants, plant viruses, and bacteria which comprise genes expressed in plant cells such as *Agrobacterium* or *Rhizobium*. Examples of promoters under developmental control include promoters that preferentially initiate transcription in certain tissues, such as leaves, roots, or seeds. Such promoters are referred to as "tissue preferred". Promoters which initiate transcription only in certain tissue are referred to as "tissue specific". A "cell type" specific promoter primarily drives expression in certain cell types in one or more organs, for example, vascular cells in roots or leaves. An "inducible" or "repressible" promoter is a promoter which is under environmental control. Examples of environmental conditions that may effect transcription by inducible promoters include anaerobic conditions or the presence of light. Tissue specific, tissue preferred, cell type specific, and inducible promoters constitute the class of "non-constitutive" promoters. A "constitutive" promoter is a promoter which is active under most environmental conditions.

As used herein "recombinant" includes reference to a cell or vector, that has been modified by the introduction of a heterologous nucleic acid or that the cell is derived from a cell so modified. Thus, for example, recombinant cells express genes that are not found

in identical form within the native (non-recombinant) form of the cell or express native genes that are otherwise abnormally expressed, under-expressed or not expressed at all as a result of human intervention. The term "recombinant" as used herein does not encompass the alteration of the cell or vector by naturally occurring events (e.g., spontaneous mutation, natural transformation/transduction/transposition) such as those occurring without human intervention.

As used herein, a "recombinant expression cassette" is a nucleic acid construct, generated recombinantly or synthetically, with a series of specified nucleic acid elements which permit transcription of a particular nucleic acid in a host cell. The recombinant expression cassette can be incorporated into a plasmid, chromosome, mitochondrial DNA, plastid DNA, virus, or nucleic acid fragment. Typically, the recombinant expression cassette portion of an expression vector includes, among other sequences, a nucleic acid to be transcribed, and a promoter.

The term "residue" or "amino acid residue" or "amino acid" are used interchangeably herein to refer to an amino acid that is incorporated into a protein, polypeptide, or peptide (collectively "protein"). The amino acid may be a naturally occurring amino acid and, unless otherwise limited, may encompass non-natural analogs of natural amino acids that can function in a similar manner as naturally occurring amino acids.

The term "selectively hybridizes" includes reference to hybridization, under stringent hybridization conditions, of a nucleic acid sequence to a specified nucleic acid target sequence to a detectably greater degree (e.g., at least 2-fold over background) than its hybridization to non-target nucleic acid sequences and to the substantial exclusion of non-target nucleic acids. Selectively hybridizing sequences typically have about at least 80% sequence identity, preferably 90% sequence identity, and most preferably 100% sequence identity (i.e., complementary) with each other.

The term "stringent conditions" or "stringent hybridization conditions" includes reference to conditions under which a probe will selectively hybridize to its target sequence, to a detectably greater degree than to other sequences (e.g., at least 2-fold over background). Stringent conditions are sequence-dependent and will be different in different circumstances. By controlling the stringency of the hybridization and/or washing conditions, target sequences can be identified which are 100% complementary to the probe (homologous probing). Alternatively, stringency conditions can be adjusted to allow some mismatching in sequences so that lower degrees of similarity are detected (heterologous

stringency of hybridization and/or wash solutions are inherently described. If the desired degree of mismatching results in a T_m of less than 45 °C (aqueous solution) or 32 °C (formamide solution) it is preferred to increase the SSC concentration so that a higher temperature can be used. Hybridization and/or wash conditions can be applied for at least 10, 30, 60, 90, 120, or 240 minutes. An extensive guide to the hybridization of nucleic acids is found in Tijssen, *Laboratory Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Acid Probes*, Part I, Chapter 2 "Overview of principles of hybridization and the strategy of nucleic acid probe assays", Elsevier, New York (1993); and *Current Protocols in Molecular Biology*, Chapter 2, Ausubel, *et al.*, Eds., Greene Publishing and Wiley-Interscience, New York (1995).

As used herein, "transgenic plant" includes reference to a plant which comprises within its genome a heterologous polynucleotide. Generally, the heterologous polynucleotide is stably integrated within the genome such that the polynucleotide is passed on to successive generations. The heterologous polynucleotide may be integrated into the genome alone or as part of a recombinant expression cassette. "Transgenic" is used herein to include any cell, cell line, callus, tissue, plant part or plant, the genotype of which has been altered by the presence of heterologous nucleic acid including those transgenics initially so altered as well as those created by sexual crosses or asexual propagation from the initial transgenic. The term "transgenic" as used herein does not encompass the alteration of the genome (chromosomal or extra-chromosomal) by conventional plant breeding methods or by naturally occurring events such as random cross-fertilization, non-recombinant viral infection, non-recombinant bacterial transformation, non-recombinant transposition, or spontaneous mutation.

As used herein, "vector" includes reference to a nucleic acid used in introduction of a polynucleotide of the present invention into a host cell. Vectors are often replicons. Expression vectors permit transcription of a nucleic acid inserted therein.

The following terms are used to describe the sequence relationships between a polynucleotide/polypeptide of the present invention with a reference polynucleotide/polypeptide: (a) "reference sequence", (b) "comparison window", (c) "sequence identity", and (d) "percentage of sequence identity".

(a) As used herein, "reference sequence" is a defined sequence used as a basis for sequence comparison with a polynucleotide/polypeptide of the present invention. A reference sequence may be a subset or the entirety of a specified sequence; for example, as

probing). Generally, a probe is less than about 1000 nucleotides in length, optionally less than 500 nucleotides in length.

Typically, stringent conditions will be those in which the salt concentration is less than about 1.5 M Na ion, typically about 0.01 to 1.0 M Na ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes (*e.g.*, 10 to 50 nucleotides) and at least about 60°C for long probes (*e.g.*, greater than 50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide. Exemplary low stringency conditions include hybridization with a buffer solution of 30 to 35% formamide, 1 M NaCl, 1% SDS (sodium dodecyl sulphate) at 37°C, and a wash in 1X to 2X SSC (20X SSC = 3.0 M NaCl/0.3 M trisodium citrate) at 50 to 55°C. Exemplary moderate stringency conditions include hybridization in 40 to 45% formamide, 1 M NaCl, 1% SDS at 37°C, and a wash in 0.5X to 1X SSC at 55 to 60°C. Exemplary high stringency conditions include hybridization in 50% formamide, 1 M NaCl, 1% SDS at 37°C, and a wash in 0.1X SSC at 60 to 65°C.

Specificity is typically the function of post-hybridization washes, the critical factors being the ionic strength and temperature of the final wash solution. For DNA-DNA hybrids, the T_m can be approximated from the equation of Meinkoth and Wahl, *Anal. Biochem.*, 138:267-284 (1984): $T_m = 81.5\ °C + 16.6\ (\log M) + 0.41\ (\%GC) - 0.61\ (\%\ form) - 500/L$; where M is the molarity of monovalent cations, %GC is the percentage of guanosine and cytosine nucleotides in the DNA, % form is the percentage of formamide in the hybridization solution, and L is the length of the hybrid in base pairs. The T_m is the temperature (under defined ionic strength and pH) at which 50% of a complementary target sequence hybridizes to a perfectly matched probe. T_m is reduced by about 1°C for each 1% of mismatching; thus, T_m, hybridization and/or wash conditions can be adjusted to hybridize to sequences of the desired identity. For example, if sequences with $\geq$90% identity are sought, the T_m can be decreased 10°C. Generally, stringent conditions are selected to be about 5°C lower than the thermal melting point (T_m) for the specific sequence and its complement at a defined ionic strength and pH. However, severely stringent conditions can utilize a hybridization and/or wash at 1, 2, 3, or 4 °C lower than the thermal melting point (T_m); moderately stringent conditions can utilize a hybridization and/or wash at 6, 7, 8, 9, or 10 °C lower than the thermal melting point (T_m); low stringency conditions can utilize a hybridization and/or wash at 11, 12, 13, 14, 15, or 20 °C lower than the thermal melting point (T_m). Using the equation, hybridization and wash compositions, and desired T_m, those of ordinary skill will understand that variations in the

a segment of a full-length cDNA or gene sequence, or the complete cDNA or gene sequence.

(b) As used herein, "comparison window" includes reference to a contiguous and specified segment of a polynucleotide/polypeptide sequence, wherein the polynucleotide/polypeptide sequence may be compared to a reference sequence and wherein the portion of the polynucleotide/polypeptide sequence in the comparison window may comprise additions or deletions (i.e., gaps) compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. Generally, the comparison window is at least 20 contiguous nucleotides/amino acids residues in length, and optionally can be 30, 40, 50, 100, or longer. Those of skill in the art understand that to avoid a high similarity to a reference sequence due to inclusion of gaps in the polynucleotide/polypeptide sequence, a gap penalty is typically introduced and is subtracted from the number of matches.

Methods of alignment of sequences for comparison are well-known in the art. Optimal alignment of sequences for comparison may be conducted by the local homology algorithm of Smith and Waterman, *Adv. Appl. Math.* 2: 482 (1981); by the homology alignment algorithm of Needleman and Wunsch, *J. Mol. Biol.* 48: 443 (1970); by the search for similarity method of Pearson and Lipman, *Proc. Natl. Acad. Sci.* 85: 2444 (1988); by computerized implementations of these algorithms, including, but not limited to: CLUSTAL in the PC/Gene program by Intelligenetics, Mountain View, California; GAP, BESTFIT, BLAST, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group (GCG), 575 Science Dr., Madison, Wisconsin, USA; the CLUSTAL program is well described by Higgins and Sharp, *Gene* 73: 237-244 (1988); Higgins and Sharp, *CABIOS* 5: 151-153 (1989); Corpet, *et al., Nucleic Acids Research* 16: 10881-90 (1988); Huang, *et al., Computer Applications in the Biosciences* 8: 155-65 (1992), and Pearson, *et al., Methods in Molecular Biology* 24: 307-331 (1994).

The BLAST family of programs which can be used for database similarity searches includes: BLASTN for nucleotide query sequences against nucleotide database sequences; BLASTX for nucleotide query sequences against protein database sequences; BLASTP for protein query sequences against protein database sequences; TBLASTN for protein query sequences against nucleotide database sequences; and TBLASTX for nucleotide query sequences against nucleotide database sequences. See, *Current Protocols in Molecular Biology*, Chapter 19, Ausubel, *et al.*, Eds., Greene Publishing and Wiley-Interscience, New

York (1995); Altschul *et al.*, *J. Mol. Biol.*, 215:403-410 (1990); and, Altschul *et al.*, *Nucleic Acids Res.* 25:3389-3402 (1997).

Software for performing BLAST analyses is publicly available, e.g., through the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/). This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold. These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are then extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always > 0) and N (penalty score for mismatching residues; always < 0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) of 10, a cutoff of 100, M=5, N=-4, and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength (W) of 3, an expectation (E) of 10, and the BLOSUM62 scoring matrix (*see* Henikoff & Henikoff (1989) *Proc. Natl. Acad. Sci. USA* 89:10915).

In addition to calculating percent sequence identity, the BLAST algorithm also performs a statistical analysis of the similarity between two sequences (*see, e.g.*, Karlin & Altschul, *Proc. Nat'l. Acad. Sci. USA* 90:5873-5877 (1993)). One measure of similarity provided by the BLAST algorithm is the smallest sum probability (P(N)), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance.

BLAST searches assume that proteins can be modeled as random sequences. However, many real proteins comprise regions of nonrandom sequences which may be homopolymeric tracts, short-period repeats, or regions enriched in one or more amino acids. Such low-complexity regions may be aligned between unrelated proteins even

though other regions of the protein are entirely dissimilar. A number of low-complexity filter programs can be employed to reduce such low-complexity alignments. For example, the SEG (Wooten and Federhen, *Comput. Chem.*, 17:149-163 (1993)) and XNU (Claverie and States, *Comput. Chem.*, 17:191-201 (1993)) low-complexity filters can be employed alone or in combination.

Unless otherwise stated, nucleotide and protein identity/similarity values provided herein are calculated using GAP (GCG Version 10) under default values.

GAP (Global Alignment Program) can also be used to compare a polynucleotide or polypeptide of the present invention with a reference sequence. GAP uses the algorithm of Needleman and Wunsch (*J. Mol. Biol.* 48: 443-453, 1970) to find the alignment of two complete sequences that maximizes the number of matches and minimizes the number of gaps. GAP considers all possible alignments and gap positions and creates the alignment with the largest number of matched bases and the fewest gaps. It allows for the provision of a gap creation penalty and a gap extension penalty in units of matched bases. GAP must make a profit of gap creation penalty number of matches for each gap it inserts. If a gap extension penalty greater than zero is chosen, GAP must, in addition, make a profit for each gap inserted of the length of the gap times the gap extension penalty. Default gap creation penalty values and gap extension penalty values in Version 10 of the Wisconsin Genetics Software Package for protein sequences are 8 and 2, respectively. For nucleotide sequences the default gap creation penalty is 50 while the default gap extension penalty is 3. The gap creation and gap extension penalties can be expressed as an integer selected from the group of integers consisting of from 0 to 100. Thus, for example, the gap creation and gap extension penalties can each independently be: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 30, 40, 50, 60 or greater.

GAP presents one member of the family of best alignments. There may be many members of this family, but no other member has a better quality. GAP displays four figures of merit for alignments: Quality, Ratio, Identity, and Similarity. The Quality is the metric maximized in order to align the sequences. Ratio is the quality divided by the number of bases in the shorter segment. Percent Identity is the percent of the symbols that actually match. Percent Similarity is the percent of the symbols that are similar. Symbols that are across from gaps are ignored. A similarity is scored when the scoring matrix value for a pair of symbols is greater than or equal to 0.50, the similarity threshold. The scoring matrix used in Version 10 of the Wisconsin Genetics Software Package is BLOSUM62 (*see* Henikoff & Henikoff (1989) *Proc. Natl. Acad. Sci. USA* 89:10915).

Multiple alignment of the sequences can be performed using the CLUSTAL method of alignment (Higgins and Sharp (1989) *CABIOS.* 5:151-153) with the default parameters (GAP PENALTY=10, GAP LENGTH PENALTY=10). Default parameters for pairwise alignments using the CLUSTAL method are KTUPLE 1, GAP PENALTY=3, WINDOW=5 and DIAGONALS SAVED=5.

(c) As used herein, "sequence identity" or "identity" in the context of two nucleic acid or polypeptide sequences includes reference to the residues in the two sequences which are the same when aligned for maximum correspondence over a specified comparison window. When percentage of sequence identity is used in reference to proteins it is recognized that residue positions which are not identical often differ by conservative amino acid substitutions, where amino acid residues are substituted for other amino acid residues with similar chemical properties (e.g. charge or hydrophobicity) and therefore do not change the functional properties of the molecule. Where sequences differ in conservative substitutions, the percent sequence identity may be adjusted upwards to correct for the conservative nature of the substitution. Sequences which differ by such conservative substitutions are said to have "sequence similarity" or "similarity". Means for making this adjustment are well-known to those of skill in the art. Typically this involves scoring a conservative substitution as a partial rather than a full mismatch, thereby increasing the percentage sequence identity. Thus, for example, where an identical amino acid is given a score of 1 and a non-conservative substitution is given a score of zero, a conservative substitution is given a score between zero and 1. The scoring of conservative substitutions is calculated, *e.g.*, according to the algorithm of Meyers and Miller, *Computer Applic. Biol. Sci.*, 4: 11-17 (1988) *e.g.*, as implemented in the program PC/GENE (Intelligenetics, Mountain View, California, USA).

(d) As used herein, "percentage of sequence identity" means the value determined by comparing two optimally aligned sequences over a comparison window, wherein the portion of the polynucleotide sequence in the comparison window may comprise additions or deletions (i.e., gaps) as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid base or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison and multiplying the result by 100 to yield the percentage of sequence identity.

Utilities

The present invention provides, among other things, compositions and methods for modulating (i.e., increasing or decreasing) the level of polynucleotides and polypeptides of the present invention in plants. In particular, the polynucleotides and polypeptides of the present invention can be expressed temporally or spatially, e.g., at developmental stages, in tissues, and/or in quantities, which are uncharacteristic of non-recombinantly engineered plants.

The present invention also provides isolated nucleic acids comprising polynucleotides of sufficient length and complementarity to a polynucleotide of the present invention to use as probes or amplification primers in the detection, quantitation, or isolation of gene transcripts. For example, isolated nucleic acids of the present invention can be used as probes in detecting deficiencies in the level of mRNA in screenings for desired transgenic plants, for detecting mutations in the gene (e.g., substitutions, deletions, or additions), for monitoring upregulation of expression or changes in enzyme activity in screening assays of compounds, for detection of any number of allelic variants (polymorphisms), orthologs, or paralogs of the gene, or for site directed mutagenesis in eukaryotic cells (see, e.g., U.S. Patent No. 5,565,350). The isolated nucleic acids of the present invention can also be used for recombinant expression of their encoded polypeptides, or for use as immunogens in the preparation and/or screening of antibodies. The isolated nucleic acids of the present invention can also be employed for use in sense or antisense suppression of one or more genes of the present invention in a host cell, tissue, or plant. Attachment of chemical agents which bind, intercalate, cleave and/or crosslink to the isolated nucleic acids of the present invention can also be used to modulate transcription or translation.

The present invention also provides isolated proteins comprising a polypeptide of the present invention (e.g., preproenzyme, proenzyme, or enzymes). The present invention also provides proteins comprising at least one epitope from a polypeptide of the present invention. The proteins of the present invention can be employed in assays for enzyme agonists or antagonists of enzyme function, or for use as immunogens or antigens to obtain antibodies specifically immunoreactive with a protein of the present invention. Such antibodies can be used in assays for expression levels, for identifying and/or isolating nucleic acids of the present invention from expression libraries, for identification of

homologous polypeptides from other species, or for purification of polypeptides of the present invention.

The isolated nucleic acids and polypeptides of the present invention can be used over a broad range of plant types, particularly monocots such as the species of the family *Gramineae* including *Hordeum, Secale, Oryza, Triticum, Sorghum* (e.g., *S. bicolor*) and *Zea* (e.g., *Z. mays*), and dicots such as *Glycine*.

The isolated nucleic acid and proteins of the present invention can also be used in species from the genera: *Cucurbita, Rosa, Vitis, Juglans, Fragaria, Lotus, Medicago, Onobrychis, Trifolium, Trigonella, Vigna, Citrus, Linum, Geranium, Manihot, Daucus, Arabidopsis, Brassica, Raphanus, Sinapis, Atropa, Capsicum, Datura, Hyoscyamus, Lycopersicon, Nicotiana, Solanum, Petunia, Digitalis, Majorana, Ciahorium, Helianthus, Lactuca, Bromus, Asparagus, Antirrhinum, Heterocallis, Nemesis, Pelargonium, Panieum, Pennisetum, Ranunculus, Senecio, Salpiglossis, Cucumis, Browallia, Pisum, Phaseolus, Lolium,* and *Avena*.

Nucleic Acids

The present invention provides, among other things, isolated nucleic acids of RNA, DNA, and analogs and/or chimeras thereof, comprising a polynucleotide of the present invention.

A polynucleotide of the present invention is inclusive of those in the accompanying Sequence Listing and:

(a) an isolated polynucleotide encoding a polypeptide of the present invention, such as those polypeptides included in the accompanying Sequence Listing, including exemplary polynucleotides of the present invention;

(b) an isolated polynucleotide which is the product of amplification from a plant nucleic acid library using primer pairs which selectively hybridize under stringent conditions to loci within a polynucleotide of the present invention;

(c) an isolated polynucleotide which selectively hybridizes to a polynucleotide of (a) or (b);

(d) an isolated polynucleotide having a specified sequence identity with polynucleotides of (a), (b), or (c);

(e) an isolated polynucleotide encoding a protein having a specified number of contiguous amino acids from a prototype polypeptide, wherein the protein is specifically recognized by antisera elicited by presentation of the protein and wherein the protein does

not detectably immunoreact to antisera which has been fully immunosorbed with the protein;

(f) complementary sequences of polynucleotides of (a), (b), (c), (d), or (e); and

(g) an isolated polynucleotide comprising at least a specific number of contiguous nucleotides from a polynucleotide of (a), (b), (c), (d), (e), or (f);

(h) an isolated polynucleotide from a full-length enriched cDNA library having the physico-chemical property of selectively hybridizing to a polynucleotide of (a), (b), (c), (d), (e), (f), or (g);

(i) an isolated polynucleotide made by the process of: 1) providing a full-length enriched nucleic acid library, 2) selectively hybridizing the polynucleotide to a polynucleotide of (a), (b), (c), (d), (e), (f), (g), or (h), thereby isolating the polynucleotide from the nucleic acid library.

A. Polynucleotides Encoding A Polypeptide of the Present Invention

As indicated in (a), above, the present invention provides isolated nucleic acids comprising a polynucleotide of the present invention, wherein the polynucleotide encodes a polypeptide of the present invention. Every nucleic acid sequence herein that encodes a polypeptide also, by reference to the genetic code, describes every possible silent variation of the nucleic acid. One of ordinary skill will recognize that each codon in a nucleic acid (except AUG, which is ordinarily the only codon for methionine; and UGG , which is ordinarily the only codon for tryptophan) can be modified to yield a functionally identical molecule. Thus, each silent variation of a nucleic acid which encodes a polypeptide of the present invention is implicit in each described polypeptide sequence and is within the scope of the present invention. Accordingly, the present invention includes polynucleotides of the present invention and polynucleotides encoding a polypeptide of the present invention.

B. Polynucleotides Amplified from a Plant Nucleic Acid Library

As indicated in (b), above, the present invention provides an isolated nucleic acid comprising a polynucleotide of the present invention, wherein the polynucleotides are amplified, under nucleic acid amplification conditions, from a plant nucleic acid library. Nucleic acid amplification conditions for each of the variety of amplification methods are well known to those of ordinary skill in the art. The plant nucleic acid library can be constructed from a monocot such as a cereal crop. Exemplary cereals include maize,

sorghum, alfalfa, canola, wheat, or rice. The plant nucleic acid library can also be constructed from a dicot such as soybean. *Zea mays* lines B73, PHRE1, A632, BMS-P2#10, W23, and Mo17 are known and publicly available. Other publicly known and available maize lines can be obtained from the Maize Genetics Cooperation (Urbana, IL). Wheat lines are available from the Wheat Genetics Resource Center (Manhattan, KS).

The nucleic acid library may be a cDNA library, a genomic library, or a library generally constructed from nuclear transcripts at any stage of intron processing. cDNA libraries can be normalized to increase the representation of relatively rare cDNAs. In optional embodiments, the cDNA library is constructed using an enriched full-length cDNA synthesis method. Examples of such methods include Oligo-Capping (Maruyama, K. and Sugano, S. *Gene* 138: 171-174, 1994), Biotinylated CAP Trapper (Carninci, *et al. Genomics* 37: 327-336, 1996), and CAP Retention Procedure (Edery, E., Chu, L.L., *et al. Molecular and Cellular Biology* 15: 3363-3371, 1995). Rapidly growing tissues or rapidly dividing cells are preferred for use as an mRNA source for construction of a cDNA library. Growth stages of maize are described in "How a Corn Plant Develops," Special Report No. 48, Iowa State University of Science and Technology Cooperative Extension Service, Ames, Iowa, Reprinted February 1993.

A polynucleotide of this embodiment (or subsequences thereof) can be obtained, for example, by using amplification primers which are selectively hybridized and primer extended, under nucleic acid amplification conditions, to at least two sites within a polynucleotide of the present invention, or to two sites within the nucleic acid which flank and comprise a polynucleotide of the present invention, or to a site within a polynucleotide of the present invention and a site within the nucleic acid which comprises it. Methods for obtaining 5' and/or 3' ends of a vector insert are well known in the art. See, e.g., RACE (Rapid Amplification of Complementary Ends) as described in Frohman, M. A., in PCR Protocols: A Guide to Methods and Applications, M. A. Innis, D. H. Gelfand, J. J. Sninsky, T. J. White, Eds. (Academic Press, Inc., San Diego), pp. 28-38 (1990)); see also, U.S. Pat. No. 5,470,722, and *Current Protocols in Molecular Biology*, Unit 15.6, Ausubel, *et al.*, Eds., Greene Publishing and Wiley-Interscience, New York (1995); Frohman and Martin, *Techniques* 1:165 (1989).

Optionally, the primers are complementary to a subsequence of the target nucleic acid which they amplify but may have a sequence identity ranging from about 85% to 99% relative to the polynucleotide sequence which they are designed to anneal to. As those skilled in the art will appreciate, the sites to which the primer pairs will selectively

hybridize are chosen such that a single contiguous nucleic acid can be formed under the desired nucleic acid amplification conditions. The primer length in nucleotides is selected from the group of integers consisting of from at least 15 to 50. Thus, the primers can be at least 15, 18, 20, 25, 30, 40, or 50 nucleotides in length. Those of skill will recognize that a lengthened primer sequence can be employed to increase specificity of binding (i.e., annealing) to a target sequence. A non-annealing sequence at the 5'end of a primer (a "tail") can be added, for example, to introduce a cloning site at the terminal ends of the amplicon.

The amplification products can be translated using expression systems well known to those of skill in the art. The resulting translation products can be confirmed as polypeptides of the present invention by, for example, assaying for the appropriate catalytic activity (e.g., specific activity and/or substrate specificity), or verifying the presence of one or more epitopes which are specific to a polypeptide of the present invention. Methods for protein synthesis from PCR derived templates are known in the art and available commercially. See, e.g., Amersham Life Sciences, Inc, Catalog '97, p.354.

C. Polynucleotides Which Selectively Hybridize to a Polynucleotide of (A) or (B)

As indicated in (c), above, the present invention provides isolated nucleic acids comprising polynucleotides of the present invention, wherein the polynucleotides selectively hybridize, under selective hybridization conditions, to a polynucleotide of sections (A) or (B) as discussed above. Thus, the polynucleotides of this embodiment can be used for isolating, detecting, and/or quantifying nucleic acids comprising the polynucleotides of (A) or (B). For example, polynucleotides of the present invention can be used to identify, isolate, or amplify partial or full-length clones in a deposited library. In some embodiments, the polynucleotides are genomic or cDNA sequences isolated or otherwise complementary to a cDNA from a dicot or monocot nucleic acid library. Exemplary species of monocots and dicots include, but are not limited to: maize, canola, soybean, cotton, wheat, sorghum, sunflower, alfalfa, oats, sugar cane, millet, barley, and rice. The cDNA library comprises at least 50% to 95% full-length sequences (for example, at least 50%, 60%, 70%, 80%, 90%, or 95% full-length sequences). The cDNA libraries can be normalized to increase the representation of rare sequences. See, e.g., U.S. Patent No. 5,482,845. Low stringency hybridization conditions are typically, but not exclusively, employed with sequences having a reduced sequence identity relative to complementary sequences. Moderate and high stringency conditions can optionally be

employed for sequences of greater identity. Low stringency conditions allow selective hybridization of sequences having about 70% to 80% sequence identity and can be employed to identify orthologous or paralogous sequences.

D. Polynucleotides Having a Specific Sequence Identity with the Polynucleotides of (A), (B) or (C)

As indicated in (d), above, the present invention provides isolated nucleic acids comprising polynucleotides of the present invention, wherein the polynucleotides have a specified identity at the nucleotide level to a polynucleotide as disclosed above in sections (A), (B), or (C), above. Identity can be calculated using, for example, the BLAST, CLUSTALW, or GAP algorithms under default conditions. The percentage of identity to a reference sequence is at least 50% and, rounded upwards to the nearest integer, can be expressed as an integer selected from the group of integers consisting of from 50 to 99. Thus, for example, the percentage of identity to a reference sequence can be at least 60%, 70%, 75%, 80%, 85%, 90%, or 95%.

Optionally, the polynucleotides of this embodiment will encode a polypeptide that will share an epitope with a polypeptide encoded by the polynucleotides of sections (A), (B), or (C). Thus, these polynucleotides encode a first polypeptide which elicits production of antisera comprising antibodies which are specifically reactive to a second polypeptide encoded by a polynucleotide of (A), (B), or (C). However, the first polypeptide does not bind to antisera raised against itself when the antisera has been fully immunosorbed with the first polypeptide. Hence, the polynucleotides of this embodiment can be used to generate antibodies for use in, for example, the screening of expression libraries for nucleic acids comprising polynucleotides of (A), (B), or (C), or for purification of, or in immunoassays for, polypeptides encoded by the polynucleotides of (A), (B), or (C). The polynucleotides of this embodiment comprise nucleic acid sequences which can be employed for selective hybridization to a polynucleotide encoding a polypeptide of the present invention.

Screening polypeptides for specific binding to antisera can be conveniently achieved using peptide display libraries. This method involves the screening of large collections of peptides for individual members having the desired function or structure. Antibody screening of peptide display libraries is well known in the art. The displayed peptide sequences can be from 3 to 5000 or more amino acids in length, frequently from 5-100 amino acids long, and often from about 8 to 15 amino acids long. In addition to direct

chemical synthetic methods for generating peptide libraries, several recombinant DNA methods have been described. One type involves the display of a peptide sequence on the surface of a bacteriophage or cell. Each bacteriophage or cell contains the nucleotide sequence encoding the particular displayed peptide sequence. Such methods are described in PCT patent publication Nos. 91/17271, 91/18980, 91/19818, and 93/08278. Other systems for generating libraries of peptides have aspects of both *in vitro* chemical synthesis and recombinant methods. See, PCT Patent publication Nos. 92/05258, 92/14843, and 97/20078. See also, U.S. Patent Nos. 5,658,754; and 5,643,768. Peptide display libraries, vectors, and screening kits are commercially available from such suppliers as Invitrogen (Carlsbad, CA).

E. Polynucleotides Encoding a Protein Having a Subsequence from a Prototype Polypeptide and Cross-Reactive to the Prototype Polypeptide

As indicated in (e), above, the present invention provides isolated nucleic acids comprising polynucleotides of the present invention, wherein the polynucleotides encode a protein having a subsequence of contiguous amino acids from a prototype polypeptide of the present invention such as are provided in (a), above. The length of contiguous amino acids from the prototype polypeptide is selected from the group of integers consisting of from at least 10 to the number of amino acids within the prototype sequence. Thus, for example, the polynucleotide can encode a polypeptide having a subsequence having at least 10, 15, 20, 25, 30, 35, 40, 45, or 50, contiguous amino acids from the prototype polypeptide. Further, the number of such subsequences encoded by a polynucleotide of the instant embodiment can be any integer selected from the group consisting of from 1 to 20, such as 2, 3, 4, or 5. The subsequences can be separated by any integer of nucleotides from 1 to the number of nucleotides in the sequence such as at least 5, 10, 15, 25, 50, 100, or 200 nucleotides.

The proteins encoded by polynucleotides of this embodiment, when presented as an immunogen, elicit the production of polyclonal antibodies which specifically bind to a prototype polypeptide such as but not limited to, a polypeptide encoded by the polynucleotide of (a) or (b), above. Generally, however, a protein encoded by a polynucleotide of this embodiment does not bind to antisera raised against the prototype polypeptide when the antisera has been fully immunosorbed with the prototype polypeptide. Methods of making and assaying for antibody binding specificity/affinity are well known in the art. Exemplary immunoassay formats include ELISA, competitive

immunoassays, radioimmunoassays, Western blots, indirect immunofluorescent assays and the like.

In a preferred assay method, fully immunosorbed and pooled antisera which is elicited to the prototype polypeptide can be used in a competitive binding assay to test the protein. The concentration of the prototype polypeptide required to inhibit 50% of the binding of the antisera to the prototype polypeptide is determined. If the amount of the protein required to inhibit binding is less than twice the amount of the prototype protein, then the protein is said to specifically bind to the antisera elicited to the immunogen. Accordingly, the proteins of the present invention embrace allelic variants, conservatively modified variants, and minor recombinant modifications to a prototype polypeptide.

A polynucleotide of the present invention optionally encodes a protein having a molecular weight as the non-glycosylated protein within 20% of the molecular weight of the full-length non-glycosylated polypeptides of the present invention. Molecular weight can be readily determined by SDS-PAGE under reducing conditions. Optionally, the molecular weight is within 15% of a full length polypeptide of the present invention, more preferably within 10% or 5%, and most preferably within 3%, 2%, or 1% of a full length polypeptide of the present invention.

Optionally, the polynucleotides of this embodiment will encode a protein having a specific enzymatic activity at least 50%, 60%, 80%, or 90% of a cellular extract comprising the native, endogenous full-length polypeptide of the present invention. Further, the proteins encoded by polynucleotides of this embodiment will optionally have a substantially similar affinity constant (K_m) and/or catalytic activity (i.e., the microscopic rate constant, k_{cat}) as the native endogenous, full-length protein. Those of skill in the art will recognize that k_{cat}/K_m value determines the specificity for competing substrates and is often referred to as the specificity constant. Proteins of this embodiment can have a k_{cat}/K_m value at least 10% of a full-length polypeptide of the present invention as determined using the endogenous substrate of that polypeptide. Optionally, the k_{cat}/K_m value will be at least 20%, 30%, 40%, 50%, and most preferably at least 60%, 70%, 80%, 90%, or 95% the k_{cat}/K_m value of the full-length polypeptide of the present invention. Determination of k_{cat}, K_m , and k_{cat}/K_m can be determined by any number of means well known to those of skill in the art. For example, the initial rates (i.e., the first 5% or less of the reaction) can be determined using rapid mixing and sampling techniques (e.g., continuous-flow, stopped-flow, or rapid quenching techniques), flash photolysis, or relaxation methods (e.g., temperature jumps) in conjunction with such exemplary methods

of measuring as spectrophotometry, spectrofluorimetry, nuclear magnetic resonance, or radioactive procedures. Kinetic values are conveniently obtained using a Lineweaver-Burk or Eadie-Hofstee plot.

F. Polynucleotides Complementary to the Polynucleotides of (A)-(E)

As indicated in (f), above, the present invention provides isolated nucleic acids comprising polynucleotides complementary to the polynucleotides of paragraphs A-E, above. As those of skill in the art will recognize, complementary sequences base-pair throughout the entirety of their length with the polynucleotides of sections (A)-(E) (i.e., have 100% sequence identity over their entire length). Complementary bases associate through hydrogen bonding in double stranded nucleic acids. For example, the following base pairs are complementary: guanine and cytosine; adenine and thymine; and adenine and uracil.

G. Polynucleotides Which are Subsequences of the Polynucleotides of (A)-(F)

As indicated in (g), above, the present invention provides isolated nucleic acids comprising polynucleotides which comprise at least 15 contiguous bases from the polynucleotides of sections (A) through (F) as discussed above. The length of the polynucleotide is given as an integer selected from the group consisting of from at least 15 to the length of the nucleic acid sequence from which the polynucleotide is a subsequence of. Thus, for example, polynucleotides of the present invention are inclusive of polynucleotides comprising at least 15, 20, 25, 30, 40, 50, 60, 75, or 100 contiguous nucleotides in length from the polynucleotides of (A)-(F). Optionally, the number of such subsequences encoded by a polynucleotide of the instant embodiment can be any integer selected from the group consisting of from 1 to 20, such as 2, 3, 4, or 5. The subsequences can be separated by any integer of nucleotides from 1 to the number of nucleotides in the sequence such as at least 5, 10, 15, 25, 50, 100, or 200 nucleotides.

Subsequences can be made by *in vitro* synthetic, *in vitro* biosynthetic, or *in vivo* recombinant methods. In optional embodiments, subsequences can be made by nucleic acid amplification. For example, nucleic acid primers will be constructed to selectively hybridize to a sequence (or its complement) within, or co-extensive with, the coding region.

The subsequences of the present invention can comprise structural characteristics of the sequence from which it is derived. Alternatively, the subsequences can lack certain

structural characteristics of the larger sequence from which it is derived such as a poly (A) tail. Optionally, a subsequence from a polynucleotide encoding a polypeptide having at least one epitope in common with a prototype polypeptide sequence as provided in (a), above, may encode an epitope in common with the prototype sequence. Alternatively, the subsequence may not encode an epitope in common with the prototype sequence but can be used to isolate the larger sequence by. for example, nucleic acid hybridization with the sequence from which it's derived. Subsequences can be used to modulate or detect gene expression by introducing into the subsequences compounds which bind, intercalate, cleave and/or crosslink to nucleic acids. Exemplary compounds include acridine, psoralen, phenanthroline, naphthoquinone, daunomycin or chloroethylaminoaryl conjugates.

H. Polynucleotides From a Full-length Enriched cDNA Library Having the Physico-Chemical Property of Selectively Hybridizing to a Polynucleotide of (A)-(G)

As indicated in (h), above, the present invention provides an isolated polynucleotide from a full-length enriched cDNA library having the physico-chemical property of selectively hybridizing to a polynucleotide of paragraphs (A), (B), (C), (D), (E), (F), or (G) as discussed above. Methods of constructing full-length enriched cDNA libraries are known in the art and discussed briefly below. The cDNA library comprises at least 50% to 95% full-length sequences (for example, at least 50%, 60%, 70%, 80%, 90%, or 95% full-length sequences). The cDNA library can be constructed from a variety of tissues from a monocot or dicot at a variety of developmental stages. Exemplary species include maize, wheat, rice, canola, soybean, cotton, sorghum, sunflower, alfalfa, oats, sugar cane, millet, barley, and rice. Methods of selectively hybridizing, under selective hybridization conditions, a polynucleotide from a full-length enriched library to a polynucleotide of the present invention are known to those of ordinary skill in the art. Any number of stringency conditions can be employed to allow for selective hybridization. In optional embodiments, the stringency allows for selective hybridization of sequences having at least 70%, 75%, 80%, 85%, 90%, 95%, or 98% sequence identity over the length of the hybridized region. Full-length enriched cDNA libraries can be normalized to increase the representation of rare sequences.

I. Polynucleotide Products Made by a cDNA Isolation Process

As indicated in (I), above, the present invention provides an isolated polynucleotide made by the process of: 1) providing a full-length enriched nucleic acid library, 2)

selectively hybridizing the polynucleotide to a polynucleotide of paragraphs (A), (B), (C), (D), (E), (F), (G, or (H) as discussed above, and thereby isolating the polynucleotide from the nucleic acid library. Full-length enriched nucleic acid libraries are constructed as discussed in paragraph (G) and below. Selective hybridization conditions are as discussed in paragraph (G). Nucleic acid purification procedures are well known in the art. Purification can be conveniently accomplished using solid-phase methods; such methods are well known to those of skill in the art and kits are available from commercial suppliers such as Advanced Biotechnologies (Surrey, UK). For example, a polynucleotide of paragraphs (A)-(H) can be immobilized to a solid support such as a membrane, bead, or particle. See, e.g., U.S. Patent No. 5,667,976. The polynucleotide product of the present process is selectively hybridized to an immobilized polynucleotide and the solid support is subsequently isolated from non-hybridized polynucleotides by methods including, but not limited to, centrifugation, magnetic separation, filtration, electrophoresis, and the like.

Construction of Nucleic Acids

The isolated nucleic acids of the present invention can be made using (a) standard recombinant methods, (b) synthetic techniques, or combinations thereof. In some embodiments, the polynucleotides of the present invention will be cloned, amplified, or otherwise constructed from a monocot such as maize, rice, or wheat, or a dicot such as soybean.

The nucleic acids may conveniently comprise sequences in addition to a polynucleotide of the present invention. For example, a multi-cloning site comprising one or more endonuclease restriction sites may be inserted into the nucleic acid to aid in isolation of the polynucleotide. Also, translatable sequences may be inserted to aid in the isolation of the translated polynucleotide of the present invention. For example, a hexa-histidine marker sequence provides a convenient means to purify the proteins of the present invention. A polynucleotide of the present invention can be attached to a vector, adapter, or linker for cloning and/or expression of a polynucleotide of the present invention. Additional sequences may be added to such cloning and/or expression sequences to optimize their function in cloning and/or expression, to aid in isolation of the polynucleotide, or to improve the introduction of the polynucleotide into a cell. Typically, the length of a nucleic acid of the present invention less the length of its polynucleotide of the present invention is less than 20 kilobase pairs, often less than 15 kb, and frequently less than 10 kb. Use of cloning vectors, expression vectors, adapters, and linkers is well

known and extensively described in the art. For a description of various nucleic acids see, for example, Stratagene Cloning Systems, Catalogs 1999 (La Jolla, CA); and, Amersham Life Sciences, Inc, Catalog '99 (Arlington Heights, IL).

A. Recombinant Methods for Constructing Nucleic Acids

The isolated nucleic acid compositions of this invention, such as RNA, cDNA, genomic DNA, or a hybrid thereof, can be obtained from plant biological sources using any number of cloning methodologies known to those of skill in the art. In some embodiments, oligonucleotide probes which selectively hybridize, under stringent conditions, to the polynucleotides of the present invention are used to identify the desired sequence in a cDNA or genomic DNA library. Isolation of RNA, and construction of cDNA and genomic libraries is well known to those of ordinary skill in the art. See, e.g., *Plant Molecular Biology: A Laboratory Manual*, Clark, Ed., Springer-Verlag, Berlin (1997); and, *Current Protocols in Molecular Biology*, Ausubel, *et al.*, Eds., Greene Publishing and Wiley-Interscience, New York (1995).

A1. Full-length Enriched cDNA Libraries

A number of cDNA synthesis protocols have been described which provide enriched full-length cDNA libraries. Enriched full-length cDNA libraries are constructed to comprise at least 600%, and more preferably at least 70%, 80%, 90% or 95% full-length inserts amongst clones containing inserts. The length of insert in such libraries can be at least 2,3, 4, 5, 6, 7, 8, 9, 10 or more kilobase pairs. Vectors to accommodate inserts of these sizes are known in the art and available commercially. See, e.g., Stratagene's lambda ZAP Express (cDNA cloning vector with 0 to 12 kb cloning capacity). An exemplary method of constructing a greater than 95% pure full-length cDNA library is described by Caminci *et al.*, *Genomics*, 37:327-336 (1996). Other methods for producing full-length libraries are known in the art. See, e.g., Edery *et al.*, *Mol. Cell Biol.*,15(6):3363-3371 (1995); and, PCT Application WO 96/34981.

A2 Normalized or Subtracted cDNA Libraries

A non-normalized cDNA library represents the mRNA population of the tissue it was made from. Since unique clones are out-numbered by clones derived from highly expressed genes their isolation can be laborious. Normalization of a cDNA library is the process of creating a library in which each clone is more equally represented.

Construction of normalized libraries is described in Ko, *Nucl. Acids. Res.*, 18(19):5705-5711 (1990); Patanjali *et al., Proc. Natl. Acad. U.S.A.*, 88:1943-1947 (1991); U.S. Patents 5,482,685, 5,482,845, and 5,637,685. In an exemplary method described by Soares *et al.*, normalization resulted in reduction of the abundance of clones from a range of four orders of magnitude to a narrow range of only 1 order of magnitude. *Proc. Natl. Acad. Sci. USA*, 91:9228-9232 (1994).

Subtracted cDNA libraries are another means to increase the proportion of less abundant cDNA species. In this procedure, cDNA prepared from one pool of mRNA is depleted of sequences present in a second pool of mRNA by hybridization. The cDNA:mRNA hybrids are removed and the remaining un-hybridized cDNA pool is enriched for sequences unique to that pool. See, *Foote et al.* in, *Plant Molecular Biology: A Laboratory Manual*, Clark, Ed., Springer-Verlag, Berlin (1997); Kho and Zarbl, *Technique*, 3(2):58-63 (1991); Sive and St. John, *Nucl. Acids Res.*, 16(22):10937 (1988); *Current Protocols in Molecular Biology*, Ausubel, *et al.*, Eds., Greene Publishing and Wiley-Interscience, New York (1995); and, Swaroop *et al., Nucl. Acids Res.*, 19)8):1954 (1991). cDNA subtraction kits are commercially available. See, e.g., PCR-Select (Clontech, Palo Alto, CA).

To construct genomic libraries, large segments of genomic DNA are generated by fragmentation, e.g. using restriction endonucleases, and are ligated with vector DNA to form concatemers that can be packaged into the appropriate vector. Methodologies to accomplish these ends, and sequencing methods to verify the sequence of nucleic acids are well known in the art. Examples of appropriate molecular biological techniques and instructions sufficient to direct persons of skill through many construction, cloning, and screening methodologies are found in Sambrook, *et al.*, *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory Vols. 1-3 (1989), Methods in Enzymology, Vol. 152: *Guide to Molecular Cloning Techniques*, Berger and Kimmel, Eds., San Diego: Academic Press, Inc. (1987), *Current Protocols in Molecular Biology*, Ausubel, *et al.*, Eds., Greene Publishing and Wiley-Interscience, New York (1995); *Plant Molecular Biology: A Laboratory Manual*, Clark, Ed., Springer-Verlag, Berlin (1997). Kits for construction of genomic libraries are also commercially available.

The cDNA or genomic library can be screened using a probe based upon the sequence of a polynucleotide of the present invention such as those disclosed herein. Probes may be used to hybridize with genomic DNA or cDNA sequences to isolate homologous genes in the same or different plant species. Those of skill in the art will

appreciate that various degrees of stringency of hybridization can be employed in the assay; and either the hybridization or the wash medium can be stringent.

The nucleic acids of interest can also be amplified from nucleic acid samples using amplification techniques. For instance, polymerase chain reaction (PCR) technology can be used to amplify the sequences of polynucleotides of the present invention and related genes directly from genomic DNA or cDNA libraries. PCR and other *in vitro* amplification methods may also be useful, for example, to clone nucleic acid sequences that code for proteins to be expressed, to make nucleic acids to use as probes for detecting the presence of the desired mRNA in samples, for nucleic acid sequencing, or for other purposes. The T4 gene 32 protein (Boehringer Mannheim) can be used to improve yield of long PCR products.

PCR-based screening methods have been described. Wilfinger *et al.* describe a PCR-based method in which the longest cDNA is identified in the first step so that incomplete clones can be eliminated from study. *BioTechniques*, 22(3): 481-486 (1997). Such methods are particularly effective in combination with a full-length cDNA construction methodology, above.

B. Synthetic Methods for Constructing Nucleic Acids

The isolated nucleic acids of the present invention can also be prepared by direct chemical synthesis by methods such as the phosphotriester method of Narang *et al., Meth. Enzymol.* 68: 90-99 (1979); the phosphodiester method of Brown *et al., Meth. Enzymol.* 68: 109-151 (1979); the diethylphosphoramidite method of Beaucage *et al., Tetra. Lett.* 22: 1859-1862 (1981); the solid phase phosphoramidite triester method described by Beaucage and Caruthers, *Tetra. Letts.* 22(20): 1859-1862 (1981), *e.g.*, using an automated synthesizer, *e.g.*, as described in Needham-VanDevanter *et al., Nucleic Acids Res.*, 12: 6159-6168 (1984); and, the solid support method of U.S. Patent No. 4,458,066. Chemical synthesis generally produces a single stranded oligonucleotide. This may be converted into double stranded DNA by hybridization with a complementary sequence, or by polymerization with a DNA polymerase using the single strand as a template. One of skill will recognize that while chemical synthesis of DNA is best employed for sequences of about 100 bases or less, longer sequences may be obtained by the ligation of shorter sequences.

Rec mbinant Expression Cassettes

The present invention further provides recombinant expression cassettes comprising a nucleic acid of the present invention. A nucleic acid sequence coding for the desired polypeptide of the present invention. for example a cDNA or a genomic sequence encoding a full length polypeptide of the present invention, can be used to construct a recombinant expression cassette which can be introduced into the desired host cell. A recombinant expression cassette will typically comprise a polynucleotide of the present invention operably linked to transcriptional initiation regulatory sequences which will direct the transcription of the polynucleotide in the intended host cell, such as tissues of a transformed plant.

For example, plant expression vectors may include (1) a cloned plant gene under the transcriptional control of 5' and 3' regulatory sequences and (2) a dominant selectable marker. Such plant expression vectors may also contain, if desired, a promoter regulatory region (e.g., one conferring inducible or constitutive, environmentally- or developmentally-regulated, or cell- or tissue-specific/selective expression), a transcription initiation start site, a ribosome binding site, an RNA processing signal, a transcription termination site, and/or a polyadenylation signal.

A plant promoter fragment can be employed which will direct expression of a polynucleotide of the present invention in all tissues of a regenerated plant. Such promoters are referred to herein as "constitutive" promoters and are active under most environmental conditions and states of development or cell differentiation. Examples of constitutive promoters include the cauliflower mosaic virus (CaMV) 35S transcription initiation region, the 1'- or 2'- promoter derived from T-DNA of *Agrobacterium tumefaciens*, the ubiquitin 1 promoter, the Smas promoter, the cinnamyl alcohol dehydrogenase promoter (U.S. Patent No. 5,683,439), the *Nos* promoter, the pEmu promoter, the rubisco promoter, and the GRP1-8 promoter.

Alternatively, the plant promoter can direct expression of a polynucleotide of the present invention in a specific tissue or may be otherwise under more precise environmental or developmental control. Such promoters are referred to here as "inducible" promoters. Environmental conditions that may effect transcription by inducible promoters include pathogen attack, anaerobic conditions, or the presence of light. Examples of inducible promoters are the Adh1 promoter which is inducible by hypoxia or cold stress, the Hsp70 promoter which is inducible by heat stress, and the PPDK promoter which is inducible by light.

Examples of promoters under developmental control include promoters that initiate transcription only, or preferentially, in certain tissues, such as leaves, roots, fruit, seeds, or flowers. Exemplary promoters include the anther specific promoter 5126 (U.S. Patent Nos. 5,689,049 and 5,689,051), glob-1 promoter, and gamma-zein promoter. See also US patent applications 60/163,114, filed November 2, 1999, and 60/155,859, filed September 24, 1999. The operation of a promoter may also vary depending on its location in the genome. Thus, an inducible promoter may become fully or partially constitutive in certain locations.

Both heterologous and non-heterologous (i.e., endogenous) promoters can be employed to direct expression of the nucleic acids of the present invention. These promoters can also be used, for example, in recombinant expression cassettes to drive expression of antisense nucleic acids to reduce, increase, or alter concentration and/or composition of the proteins of the present invention in a desired tissue. Thus, in some embodiments, the nucleic acid construct will comprise a promoter, functional in a plant cell, operably linked to a polynucleotide of the present invention. Promoters useful in these embodiments include the endogenous promoters driving expression of a polypeptide of the present invention.

In some embodiments, isolated nucleic acids which serve as promoter or enhancer elements can be introduced in the appropriate position (generally upstream) of a non-heterologous form of a polynucleotide of the present invention so as to up or down regulate expression of a polynucleotide of the present invention. For example, endogenous promoters can be altered *in vivo* by mutation, deletion, and/or substitution (see, Kmiec, U.S. Patent 5,565,350; Zarling *et al.*, WO 93/22443), or isolated promoters can be introduced into a plant cell in the proper orientation and distance from a cognate gene of a polynucleotide of the present invention so as to control the expression of the gene. Gene expression can be modulated under conditions suitable for plant growth so as to alter the total concentration and/or alter the composition of the polypeptides of the present invention in plant cell. Thus, the present invention provides compositions, and methods for making, heterologous promoters and/or enhancers operably linked to a native, endogenous (i.e., non-heterologous) form of a polynucleotide of the present invention.

If polypeptide expression is desired, it is generally desirable to include a polyadenylation region at the 3'-end of a polynucleotide coding region. The polyadenylation region can be derived from the natural gene, from a variety of other plant genes, or from T-DNA. The 3' end sequence to be added can be derived from, for

example, the nopaline synthase or octopine synthase genes, or alternatively from another plant gene, or less preferably from any other eukaryotic gene.

An intron sequence can be added to the 5' untranslated region or the coding sequence of the partial coding sequence to increase the amount of the mature message that accumulates in the cytosol. Inclusion of a spliceable intron in the transcription unit in both plant and animal expression constructs has been shown to increase gene expression at both the mRNA and protein levels up to 1000-fold. Buchman and Berg, *Mol. Cell Biol.* 8: 4395-4405 (1988); Callis *et al.*, *Genes Dev.* 1: 1183-1200 (1987). Such intron enhancement of gene expression is typically greatest when placed near the 5' end of the transcription unit. Use of maize introns Adh1-S intron 1, 2, and 6, the Bronze-1 intron are known in the art. See generally, *The Maize Handbook*, Chapter 116, Freeling and Walbot, Eds., Springer, New York (1994). The vector comprising the sequences from a polynucleotide of the present invention will typically comprise a marker gene which confers a selectable phenotype on plant cells. Typical vectors useful for expression of genes in higher plants are well known in the art and include vectors derived from the tumor-inducing (Ti) plasmid of *Agrobacterium tumefaciens* described by Rogers *et al.*, Meth. in Enzymol., 153:253-277 (1987).

A polynucleotide of the present invention can be expressed in either sense or anti-sense orientation as desired. It will be appreciated that control of gene expression in either sense or anti-sense orientation can have a direct impact on the observable plant characteristics. Antisense technology can be conveniently used to inhibit gene expression in plants. To accomplish this, a nucleic acid segment from the desired gene is cloned and operably linked to a promoter such that the anti-sense strand of RNA will be transcribed. The construct is then transformed into plants and the antisense strand of RNA is produced. In plant cells, it has been shown that antisense RNA inhibits gene expression by preventing the accumulation of mRNA which encodes the enzyme of interest, see, e.g., Sheehy *et al.*, *Proc. Nat'l. Acad. Sci. (USA)* 85: 8805-8809 (1988); and Hiatt *et al.*, U.S. Patent No. 4,801,340.

Another method of suppression is sense suppression (i.e., co-supression). Introduction of nucleic acid configured in the sense orientation has been shown to be an effective means by which to block the transcription of target genes. For an example of the use of this method to modulate expression of endogenous genes see, Napoli *et al.*, *The Plant Cell* 2: 279-289 (1990) and U.S. Patent No. 5,034,323.

Catalytic RNA molecules or ribozymes can also be used to inhibit expression of plant genes. It is possible to design ribozymes that specifically pair with virtually any target RNA and cleave the phosphodiester backbone at a specific location, thereby functionally inactivating the target RNA. In carrying out this cleavage, the ribozyme is not itself altered, and is thus capable of recycling and cleaving other molecules, making it a true enzyme. The inclusion of ribozyme sequences within antisense RNAs confers RNA-cleaving activity upon them, thereby increasing the activity of the constructs. The design and use of target RNA-specific ribozymes is described in Haseloff *et al.*, *Nature* 334: 585-591 (1988).

A variety of cross-linking agents, alkylating agents and radical generating species as pendant groups on polynucleotides of the present invention can be used to bind, label, detect, and/or cleave nucleic acids. For example, Vlassov, V. V., *et al.*, *Nucleic Acids Res* (1986) 14:4065-4076, describe covalent bonding of a single-stranded DNA fragment with alkylating derivatives of nucleotides complementary to target sequences. A report of similar work by the same group is that by Knorre, D. G., *et al.*, *Biochimie* (1985) 67:785-789. Iverson and Dervan also showed sequence-specific cleavage of single-stranded DNA mediated by incorporation of a modified nucleotide which was capable of activating cleavage (*J Am Chem Soc* (1987) 109:1241-1243). Meyer, R. B., *et al.*, *J Am Chem Soc* (1989) 111:8517-8519, effect covalent crosslinking to a target nucleotide using an alkylating agent complementary to the single-stranded target nucleotide sequence. A photoactivated crosslinking to single-stranded oligonucleotides mediated by psoralen was disclosed by Lee, B. L., *et al.*, *Biochemistry* (1988) 27:3197-3203. Use of crosslinking in triple-helix forming probes was also disclosed by Home, *et al.*, *J Am Chem Soc* (1990) 112:2435-2437. Use of N4, N4-ethanocytosine as an alkylating agent to crosslink to single-stranded oligonucleotides has also been described by Webb and Matteucci, *J Am Chem Soc* (1986) 108:2764-2765; *Nucleic Acids Res* (1986) 14:7661-7674; Feteritz *et al.*, *J. Am. Chem. Soc.* 113:4000 (1991). Various compounds to bind, detect, label, and/or cleave nucleic acids are known in the art. See, for example, U.S. Patent Nos. 5,543,507; 5,672,593; 5,484,908; 5,256,648; and, 5,681941.

Proteins

The isolated proteins of the present invention comprise a polypeptide having at least 10 amino acids from a polypeptide of the present invention (or conservative variants thereof) such as those encoded by any one of the polynucleotides of the present invention

as discussed more fully above. The proteins of the present invention or variants thereof can comprise any number of contiguous amino acid residues from a polypeptide of the present invention, wherein that number is selected from the group of integers consisting of from 10 to the number of residues in a full-length polypeptide of the present invention. Optionally, this subsequence of contiguous amino acids is at least 15, 20, 25, 30, 35, or 40 amino acids in length, often at least 50, 60, 70, 80, or 90 amino acids in length. Further, the number of such subsequences can be any integer selected from the group consisting of from 1 to 20, such as 2, 3, 4, or 5.

The present invention further provides a protein comprising a polypeptide having a specified sequence identity/similarity with a polypeptide of the present invention. The percentage of sequence identity/similarity is an integer selected from the group consisting of from 50 to 99. Exemplary sequence identity/similarity values include 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, and 95%. Sequence identity can be determined using, for example, the GAP, CLUSTALW, or BLAST algorithms.

As those of skill will appreciate, the present invention includes, but is not limited to, catalytically active polypeptides of the present invention (i.e., enzymes). Catalytically active polypeptides have a specific activity of at least 20%, 30%, or 40%, and preferably at least 50%, 60%, or 70%, and most preferably at least 80%, 90%, or 95% that of the native (non-synthetic), endogenous polypeptide. Further, the substrate specificity (k_{cat}/K_m) is optionally substantially similar to the native (non-synthetic), endogenous polypeptide. Typically, the K_m will be at least 30%, 40%, or 50%, that of the native (non-synthetic), endogenous polypeptide; and more preferably at least 60%, 70%, 80%, or 90%. Methods of assaying and quantifying measures of enzymatic activity and substrate specificity (k_{cat}/K_m), are well known to those of skill in the art.

Generally, the proteins of the present invention will, when presented as an immunogen, elicit production of an antibody specifically reactive to a polypeptide of the present invention. Further, the proteins of the present invention will not bind to antisera raised against a polypeptide of the present invention which has been fully immunosorbed with the same polypeptide. Immunoassays for determining binding are well known to those of skill in the art. A preferred immunoassay is a competitive immunoassay. Thus, the proteins of the present invention can be employed as immunogens for constructing antibodies immunoreactive to a protein of the present invention for such exemplary utilities as immunoassays or protein purification techniques.

Expression of Proteins in Host Cells

Using the nucleic acids of the present invention, one may express a protein of the present invention in a recombinantly engineered cell such as bacteria, yeast, insect, mammalian, or preferably plant cells. The cells produce the protein in a non-natural condition (e.g., in quantity, composition, location, and/or time), because they have been genetically altered through human intervention to do so.

It is expected that those of skill in the art are knowledgeable in the numerous expression systems available for expression of a nucleic acid encoding a protein of the present invention. No attempt to describe in detail the various methods known for the expression of proteins in prokaryotes or eukaryotes will be made.

In brief summary, the expression of isolated nucleic acids encoding a protein of the present invention will typically be achieved by operably linking, for example, the DNA or cDNA to a promoter (which is either constitutive or regulatable), followed by incorporation into an expression vector. The vectors can be suitable for replication and integration in either prokaryotes or eukaryotes. Typical expression vectors contain transcription and translation terminators, initiation sequences, and promoters useful for regulation of the expression of the DNA encoding a protein of the present invention. To obtain high level expression of a cloned gene, it is desirable to construct expression vectors which contain, at the minimum, a strong promoter to direct transcription, a ribosome binding site for translational initiation, and a transcription/translation terminator. One of skill would recognize that modifications can be made to a protein of the present invention without diminishing its biological activity. Some modifications may be made to facilitate the cloning, expression, or incorporation of the targeting molecule into a fusion protein. Such modifications are well known to those of skill in the art and include, for example, a methionine added at the amino terminus to provide an initiation site, or additional amino acids (e.g., poly His) placed on either terminus to create conveniently located purification sequences. Restriction sites or termination codons can also be introduced.

Synthesis of Proteins

The proteins of the present invention can be constructed using non-cellular synthetic methods. Solid phase synthesis of proteins of less than about 50 amino acids in length may be accomplished by attaching the C-terminal amino acid of the sequence to an insoluble support followed by sequential addition of the remaining amino acids in the sequence. Techniques for solid phase synthesis are described by Barany and Merrifield,

Solid-Phase Peptide Synthesis, pp. 3-284 in *The Peptides: Analysis, Synthesis, Biology. Vol. 2: Special Methods in Peptide Synthesis, Part A.*; Merrifield, *et al., J. Am. Chem. Soc.* 85: 2149-2156 (1963), and *Stewart et al., Solid Phase Peptide Synthesis, 2nd ed.*, Pierce Chem. Co., Rockford, Ill. (1984). Proteins of greater length may be synthesized by condensation of the amino and carboxy termini of shorter fragments. Methods of forming peptide bonds by activation of a carboxy terminal end (e.g., by the use of the coupling reagent N,N'-dicycylohexylcarbodiimide) are known to those of skill.

Purification of Proteins

The proteins of the present invention may be purified by standard techniques well known to those of skill in the art. Recombinantly produced proteins of the present invention can be directly expressed or expressed as a fusion protein. The recombinant protein is purified by a combination of cell lysis (e.g., sonication, French press) and affinity chromatography. For fusion products, subsequent digestion of the fusion protein with an appropriate proteolytic enzyme releases the desired recombinant protein.

The proteins of this invention, recombinant or synthetic, may be purified to substantial purity by standard techniques well known in the art, including detergent solubilization, selective precipitation with such substances as ammonium sulfate, column chromatography, immunopurification methods, and others. *See*, for instance, R. Scopes, *Protein Purification: Principles and Practice*, Springer-Verlag: New York (1982); Deutscher, *Guide to Protein Purification*, Academic Press (1990). For example, antibodies may be raised to the proteins as described herein. Purification from *E. coli* can be achieved following procedures described in U.S. Patent No. 4,511,503. The protein may then be isolated from cells expressing the protein and further purified by standard protein chemistry techniques as described herein. Detection of the expressed protein is achieved by methods known in the art and include, for example, radioimmunoassays, Western blotting techniques or immunoprecipitation.

Introduction of Nucleic Acids Into Host Cells

The method of introducing a nucleic acid of the present invention into a host cell is not critical to the instant invention. Transformation or transfection methods are conveniently used. Accordingly, a wide variety of methods have been developed to insert a DNA sequence into the genome of a host cell to obtain the transcription and/or

translation of the sequence to effect phenotypic changes in the organism. Thus, any method which provides for effective introduction of a nucleic acid may be employed.

A. Plant Transformation

A nucleic acid comprising a polynucleotide of the present invention is optionally introduced into a plant. Generally, the polynucleotide will first be incorporated into a recombinant expression cassette or vector. Isolated nucleic acid acids of the present invention can be introduced into plants according to techniques known in the art. Techniques for transforming a wide variety of higher plant species are well known and described in the technical, scientific, and patent literature. See, for example, Weising *et al., Ann. Rev. Genet.* 22: 421-477 (1988). For example, the DNA construct may be introduced directly into the genomic DNA of the plant cell using techniques such as electroporation, polyethylene glycol (PEG), poration, particle bombardment, silicon fiber delivery, or microinjection of plant cell protoplasts or embryogenic callus. See, e.g., Tomes, *et al.*, Direct DNA Transfer into Intact Plant Cells Via Microprojectile Bombardment. pp.197-213 in Plant Cell, Tissue and Organ Culture, Fundamental Methods. eds. O. L. Gamborg and G.C. Phillips. Springer-Verlag Berlin Heidelberg New York, 1995; see, U.S. Patent No. 5,990,387. The introduction of DNA constructs using PEG precipitation is described in Paszkowski *et al., Embo J.* 3: 2717-2722 (1984). Electroporation techniques are described in Fromm *et al., Proc. Natl. Acad. Sci. (USA)* 82: 5824 (1985). Ballistic transformation techniques are described in Klein *et al., Nature* 327: 70-73 (1987).

Agrobacterium tumefaciens-mediated transformation techniques are well described in the scientific literature. See, for example Horsch *et al., Science* 233: 496-498 (1984); Fraley *et al., Proc. Natl. Acad. Sci. (USA)* 80: 4803 (1983); and, *Plant Molecular Biology: A Laboratory Manual*, Chapter 8, Clark, Ed., Springer-Verlag, Berlin (1997). The DNA constructs may be combined with suitable T-DNA flanking regions and introduced into a conventional *Agrobacterium tumefaciens* host vector. The virulence functions of the *Agrobacterium tumefaciens* host will direct the insertion of the construct and adjacent marker into the plant cell DNA when the cell is infected by the bacteria. See, U.S. Patent No. 5,591,616. Although *Agrobacterium* is useful primarily in dicots, certain monocots can be transformed by *Agrobacterium*. For instance, *Agrobacterium* transformation of maize is described in U.S. Patent No. 5,550,318.

Other methods of transfection or transformation include (1) *Agrobacterium rhizogenes*-mediated transformation (see, e.g., Lichtenstein and Fuller In: Genetic Engineering, vol. 6, PWJ Rigby, Ed., London. Academic Press, 1987; and Lichtenstein, C. P., and Draper, J,. In: DNA Cloning, Vol. II, D. M. Glover, Ed., Oxford, IRI Press, 1985), Application PCT/US87/02512 (WO 88/02405 published Apr. 7, 1988) describes the use of *A. rhizogenes* strain A4 and its Ri plasmid along with *A. tumefaciens* vectors pARC8 or pARC16 (2) liposome-mediated DNA uptake (see, e.g., Freeman *et al., Plant Cell Physiol.* 25: 1353 (1984)), (3) the vortexing method (see, e.g., Kindle, *Proc. Natl. Acad. Sci., (USA)* 87: 1228 (1990).

DNA can also be introduced into plants by direct DNA transfer into pollen as described by Zhou *et al.*, Methods in Enzymology, 101:433 (1983); D. Hess, *Intern Rev. Cytol.*, 107:367 (1987); Luo *et al.*, *Plant Mol. Biol. Reporter*, 6:165 (1988). Expression of polypeptide coding genes can be obtained by injection of the DNA into reproductive organs of a plant as described by Pena *et al.*, *Nature*, 325.:274 (1987). DNA can also be injected directly into the cells of immature embryos and the rehydration of desiccated embryos as described by Neuhaus *et al.*, *Theor. Appl. Genet.*, 75:30 (1987); and Benbrook *et al.*, *in Proceedings Bio Expo* 1986, Butterworth, Stoneham, Mass., pp. 27-54 (1986). A variety of plant viruses that can be employed as vectors are known in the art and include cauliflower mosaic virus (CaMV), geminivirus, brome mosaic virus, and tobacco mosaic virus.

B. Transfection of Prokaryotes, Lower Eukaryotes, and Animal Cells

Animal and lower eukaryotic (e.g., yeast) host cells are competent or rendered competent for transfection by various means. There are several well-known methods of introducing DNA into animal cells. These include: calcium phosphate precipitation, fusion of the recipient cells with bacterial protoplasts containing the DNA, treatment of the recipient cells with liposomes containing the DNA, DEAE dextran, electroporation, biolistics, and micro-injection of the DNA directly into the cells. The transfected cells are cultured by means well known in the art. Kuchler, R.J., *Biochemical Methods in Cell Culture and Virology*, Dowden, Hutchinson and Ross, Inc. (1977).

Transgenic Plant Regeneration

Plant cells which directly result or are derived from the nucleic acid introduction techniques can be cultured to regenerate a whole plant which possesses the introduced

genotype. Such regeneration techniques often rely on manipulation of certain phytohormones in a tissue culture growth medium. Plants cells can be regenerated, e.g., from single cells, callus tissue or leaf discs according to standard plant tissue culture techniques. It is well known in the art that various cells, tissues, and organs from almost any plant can be successfully cultured to regenerate an entire plant. Plant regeneration from cultured protoplasts is described in Evans *et al.*, *Protoplasts Isolation and Culture*, *Handbook of Plant Cell Culture*, Macmillan Publishing Company, New York, pp. 124-176 (1983); and Binding, *Regeneration of Plants. Plant Protoplasts*, CRC Press, Boca Raton, pp. 21-73 (1985).

The regeneration of plants from either single plant protoplasts or various explants is well known in the art. See, for example, *Methods for Plant Molecular Biology*, A. Weissbach and H. Weissbach, eds., Academic Press, Inc., San Diego, Calif. (1988). This regeneration and growth process includes the steps of selection of transformant cells and shoots, rooting the transformant shoots and growth of the plantlets in soil. For maize cell culture and regeneration see generally, *The Maize Handbook*, Freeling and Walbot, Eds., Springer, New York (1994); *Corn and Corn Improvement*, 3rd edition, Sprague and Dudley Eds., American Society of Agronomy, Madison, Wisconsin (1988). For transformation and regeneration of maize see, Gordon-Kamm *et al.*, *The Plant Cell*, 2:603-618 (1990).

The regeneration of plants containing the polynucleotide of the present invention and introduced by *Agrobacterium* from leaf explants can be achieved as described by Horsch *et al.*, *Science*, 227:1229-1231 (1985). In this procedure, transformants are grown in the presence of a selection agent and in a medium that induces the regeneration of shoots in the plant species being transformed as described by Fraley *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)*, 80:4803 (1983). This procedure typically produces shoots within two to four weeks and these transformant shoots are then transferred to an appropriate root-inducing medium containing the selective agent and an antibiotic to prevent bacterial growth. Transgenic plants of the present invention may be fertile or sterile.

One of skill will recognize that after the recombinant expression cassette is stably incorporated in transgenic plants and confirmed to be operable, it can be introduced into other plants by sexual crossing. Any of a number of standard breeding techniques can be used, depending upon the species to be crossed. In vegetatively propagated crops, mature transgenic plants can be propagated by the taking of cuttings or by tissue culture techniques to produce multiple identical plants. Selection of desirable transgenics is made and new varieties are obtained and propagated vegetatively for commercial use. In seed

propagated crops, mature transgenic plants can be self crossed to produce a homozygous inbred plant. The inbred plant produces seed containing the newly introduced heterologous nucleic acid. These seeds can be grown to produce plants that would produce the selected phenotype. Parts obtained from the regenerated plant, such as flowers, seeds, leaves, branches, fruit, and the like are included in the invention, provided that these parts comprise cells comprising the isolated nucleic acid of the present invention. Progeny and variants, and mutants of the regenerated plants are also included within the scope of the invention, provided that these parts comprise the introduced nucleic acid sequences.

Transgenic plants expressing a polynucleotide of the present invention can be screened for transmission of the nucleic acid of the present invention by, for example, standard immunoblot and DNA detection techniques. Expression at the RNA level can be determined initially to identify and quantitate expression-positive plants. Standard techniques for RNA analysis can be employed and include PCR amplification assays using oligonucleotide primers designed to amplify only the heterologous RNA templates and solution hybridization assays using heterologous nucleic acid-specific probes. The RNA-positive plants can then analyzed for protein expression by Western immunoblot analysis using the specifically reactive antibodies of the present invention. In addition, *in situ* hybridization and immunocytochemistry according to standard protocols can be done using heterologous nucleic acid specific polynucleotide probes and antibodies, respectively, to localize sites of expression within transgenic tissue. Generally, a number of transgenic lines are usually screened for the incorporated nucleic acid to identify and select plants with the most appropriate expression profiles.

A preferred embodiment is a transgenic plant that is homozygous for the added heterologous nucleic acid; i.e., a transgenic plant that contains two added nucleic acid sequences, one gene at the same locus on each chromosome of a chromosome pair. A homozygous transgenic plant can be obtained by sexually mating (selfing) a heterozygous transgenic plant that contains a single added heterologous nucleic acid, germinating some of the seed produced and analyzing the resulting plants produced for altered expression of a polynucleotide of the present invention relative to a control plant (i.e., native, non-transgenic). Back-crossing to a parental plant and out-crossing with a non- transgenic plant are also contemplated.

Modulating Polypeptide Levels and/or Composition

The present invention further provides a method for modulating (i.e., increasing or decreasing) the concentration or ratio of the polypeptides of the present invention in a plant or part thereof. Modulation can be effected by increasing or decreasing the concentration and/or the ratio of the polypeptides of the present invention in a plant. The method comprises introducing into a plant cell a recombinant expression cassette comprising a polynucleotide of the present invention as described above to obtain a transgenic plant cell, culturing the transgenic plant cell under transgenic plant cell growing conditions, and inducing or repressing expression of a polynucleotide of the present invention in the transgenic plant for a time sufficient to modulate concentration and/or the ratios of the polypeptides in the transgenic plant or plant part.

In some embodiments, the concentration and/or ratios of polypeptides of the present invention in a plant may be modulated by altering, *in vivo* or *in vitro*, the promoter of a gene to up- or down-regulate gene expression. In some embodiments, the coding regions of native genes of the present invention can be altered via substitution, addition, insertion, or deletion to decrease activity of the encoded enzyme. See, e.g., Kmiec, U.S. Patent 5,565,350; Zarling *et al.*, WO 93/22443; Briggs, et al., U.S. Patent 5,962,764. And in some embodiments, an isolated nucleic acid (e.g., a vector) comprising a promoter sequence is transfected into a plant cell. Subsequently, a plant cell comprising the promoter operably linked to a polynucleotide of the present invention is selected for by means known to those of skill in the art such as, but not limited to, Southern blot, DNA sequencing, or PCR analysis using primers specific to the promoter and to the gene and detecting amplicons produced therefrom. A plant or plant part altered or modified by the foregoing embodiments is grown under plant forming conditions for a time sufficient to modulate the concentration and/or ratios of polypeptides of the present invention in the plant. Plant forming conditions are well known in the art and discussed briefly, *supra*.

In general, concentration or the ratios of the polypeptides is increased or decreased by at least 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% relative to a native control plant, plant part, or cell lacking the aforementioned recombinant expression cassette. Modulation in the present invention may occur during and/or subsequent to growth of the plant to the desired stage of development. Modulating nucleic acid expression temporally and/or in particular tissues can be controlled by employing the appropriate promoter operably linked to a polynucleotide of the present invention in, for example, sense or antisense orientation as discussed in greater detail, *supra*. Induction of expression of a polynucleotide of the present invention can also be controlled by

exogenous administration of an effective amount of inducing compound. Inducible promoters and inducing compounds which activate expression from these promoters are well known in the art. In preferred embodiments, the polypeptides of the present invention are modulated in monocots, particularly maize.

UTRs and Codon Preference

In general, translational efficiency has been found to be regulated by specific sequence elements in the 5' non-coding or untranslated region (5' UTR) of the RNA. Positive sequence motifs include translational initiation consensus sequences (Kozak, *Nucleic Acids Res.* 15:8125 (1987)) and the 7-methylguanosine cap structure (Drummond *et al., Nucleic Acids Res.* 13:7375 (1985)). Negative elements include stable intramolecular 5' UTR stem-loop structures (Muesing *et al., Cell* 48:691 (1987)) and AUG sequences or short open reading frames preceded by an appropriate AUG in the 5' UTR (Kozak, *supra, Rao et al., Mol. and Cell. Biol.* 8:284 (1988)). Accordingly, the present invention provides 5' and/or 3' untranslated regions for modulation of translation of heterologous coding sequences.

Further, the polypeptide-encoding segments of the polynucleotides of the present invention can be modified to alter codon usage. Altered codon usage can be employed to alter translational efficiency and/or to optimize the coding sequence for expression in a desired host such as to optimize the codon usage in a heterologous sequence for expression in maize. Codon usage in the coding regions of the polynucleotides of the present invention can be analyzed statistically using commercially available software packages such as "Codon Preference" available from the University of Wisconsin Genetics Computer Group (see Devereaux *et al., Nucleic Acids Res.* 12: 387-395 (1984)) or MacVector 4.1 (Eastman Kodak Co., New Haven, Conn.). Thus, the present invention provides a codon usage frequency characteristic of the coding region of at least one of the polynucleotides of the present invention. The number of polynucleotides that can be used to determine a codon usage frequency can be any integer from 1 to the number of polynucleotides of the present invention as provided herein. Optionally, the polynucleotides will be full-length sequences. An exemplary number of sequences for statistical analysis can be at least 1, 5, 10, 20, 50, or 100.

Sequence Shuffling

The present invention provides methods for sequence shuffling using polynucleotides of the present invention, and compositions resulting therefrom. Sequence shuffling is described in PCT publication No. WO 97/20078. See also, Zhang, J.- H., *et al. Proc. Natl. Acad. Sci. USA* 94:4504-4509 (1997). Generally, sequence shuffling provides a means for generating libraries of polynucleotides having a desired characteristic which can be selected or screened for. Libraries of recombinant polynucleotides are generated from a population of related sequence polynucleotides which comprise sequence regions which have substantial sequence identity and can be homologously recombined *in vitro* or *in vivo*. The population of sequence-recombined polynucleotides comprises a subpopulation of polynucleotides which possess desired or advantageous characteristics and which can be selected by a suitable selection or screening method. The characteristics can be any property or attribute capable of being selected for or detected in a screening system, and may include properties of: an encoded protein, a transcriptional element, a sequence controlling transcription, RNA processing, RNA stability, chromatin conformation, translation, or other expression property of a gene or transgene, a replicative element, a protein-binding element, or the like, such as any feature which confers a selectable or detectable property. In some embodiments, the selected characteristic will be a decreased K_m and/or increased K_{cat} over the wild-type protein as provided herein. In other embodiments, a protein or polynucleotide generated from sequence shuffling will have a ligand binding affinity greater than the non-shuffled wild-type polynucleotide. The increase in such properties can be at least 110%, 120%, 130%, 140% or at least 150% of the wild-type value.

Generic and Consensus Sequences

Polynucleotides and polypeptides of the present invention further include those having: (a) a generic sequence of at least two homologous polynucleotides or polypeptides, respectively, of the present invention; and, (b) a consensus sequence of at least three homologous polynucleotides or polypeptides, respectively, of the present invention. The generic sequence of the present invention comprises each species of polypeptide or polynucleotide embraced by the generic polypeptide or polynucleotide sequence, respectively. The individual species encompassed by a polynucleotide having an amino acid or nucleic acid consensus sequence can be used to generate antibodies or produce nucleic acid probes or primers to screen for homologs in other species, genera, families, orders, classes, phyla, or kingdoms. For example, a polynucleotide having a consensus

sequence from a gene family of *Zea mays* can be used to generate antibody or nucleic acid probes or primers to other *Gramineae* species such as wheat, rice, or sorghum. Alternatively, a polynucleotide having a consensus sequence generated from orthologous genes can be used to identify or isolate orthologs of other taxa. Typically, a polynucleotide having a consensus sequence will be at least 9, 10, 15, 20, 25, 30, or 40 amino acids in length, or 20, 30, 40, 50, 100, or 150 nucleotides in length. As those of skill in the art are aware, a conservative amino acid substitution can be used for amino acids which differ amongst aligned sequence but are from the same conservative substitution group as discussed above. .Optionally, no more than 1 or 2 conservative amino acids are substituted for each 10 amino acid length of consensus sequence.

Similar sequences used for generation of a consensus or generic sequence include any number and combination of allelic variants of the same gene, orthologous, or paralogous sequences as provided herein. Optionally, similar sequences used in generating a consensus or generic sequence are identified using the BLAST algorithm's smallest sum probability (P(N)). Various suppliers of sequence-analysis software are listed in chapter 7 of *Current Protocols in Molecular Biology*, F.M. Ausubel *et al.*, Eds., Current Protocols, a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc. (Supplement 30). A polynucleotide sequence is considered similar to a reference sequence if the smallest sum probability in a comparison of the test nucleic acid to the reference nucleic acid is less than about 0.1, more preferably less than about 0.01, or 0.001, and most preferably less than about 0.0001, or 0.00001. Similar polynucleotides can be aligned and a consensus or generic sequence generated using multiple sequence alignment software available from a number of commercial suppliers such as the Genetics Computer Group's (Madison, WI) PILEUP software, Vector NTI's (North Bethesda, MD) ALIGNX, or Genecode's (Ann Arbor, MI) SEQUENCHER. Conveniently, default parameters of such software can be used to generate consensus or generic sequences.

Machine Applications

The present invention provides machines, data structures, and processes for modeling or analyzing the polynucleotides and polypeptides of the present invention.

A. Machines: Data, Data Structures, Processes, and Functions

The present invention provides a machine having a memory comprising: 1) data representing a sequence of a polynucleotide or polypeptide of the present invention, 2) a

data structure which reflects the underlying organization and structure of the data and facilitates program access to data elements corresponding to logical sub-components of the sequence, 3) processes for effecting the use, analysis, or modeling of the sequence, and 4) optionally, a function or utility for the polynucleotide or polypeptide. Thus, the present invention provides a memory for storing data that can be accessed by a computer programmed to implement a process for effecting the use, analyses, or modeling of a sequence of a polynucleotide, with the memory comprising data representing the sequence of a polynucleotide of the present invention.

The machine of the present invention is typically a digital computer. The term "computer" includes one or several desktop or portable computers, computer workstations, servers (including intranet or internet servers), mainframes, and any integrated system comprising any of the above irrespective of whether the processing, memory, input, or output of the computer is remote or local, as well as any networking interconnecting the modules of the computer. The term "computer" is exclusive of computers of the United States Patent and Trademark Office or the European Patent Office when data representing the sequence of polypeptides or polynucleotides of the present invention is used for patentability searches.

The present invention contemplates providing as data a sequence of a polynucleotide of the present invention embodied in a computer readable medium. As those of skill in the art will be aware, the form of memory of a machine of the present invention, or the particular embodiment of the computer readable medium, are not critical elements of the invention and can take a variety of forms. The memory of such a machine includes, but is not limited to, ROM, or RAM, or computer readable media such as, but not limited to, magnetic media such as computer disks or hard drives, or media such as CD-ROMs, DVDs, and the like.

The present invention further contemplates providing a data structure that is also contained in memory. The data structure may be defined by the computer programs that define the processes (see below) or it may be defined by the programming of separate data storage and retrieval programs subroutines, or systems. Thus, the present invention provides a memory for storing a data structure that can be accessed by a computer programmed to implement a process for effecting the use, analysis, or modeling of a sequence of a polynucleotide. The memory comprises data representing a polynucleotide having the sequence of a polynucleotide of the present invention. The data is stored within memory. Further, a data structure, stored within memory, is associated with the data

reflecting the underlying organization and structure of the data to facilitate program access to data elements corresponding to logical sub-components of the sequence. The data structure enables the polynucleotide to be identified and manipulated by such programs.

In a further embodiment, the present invention provides a data structure that contains data representing a sequence of a polynucleotide of the present invention stored within a computer readable medium. The data structure is organized to reflect the logical structuring of the sequence, so that the sequence is easily analyzed by software programs capable of accessing the data structure. In particular, the data structures of the present invention organize the reference sequences of the present invention in a manner which allows software tools to perform a wide variety of analyses using logical elements and sub-elements of each sequence.

An example of such a data structure resembles a layered hash table, where in one dimension the base content of the sequence is represented by a string of elements A, T, C, G and N. The direction from the 5' end to the 3' end is reflected by the order from the position 0 to the position of the length of the string minus one. Such a string, corresponding to a nucleotide sequence of interest, has a certain number of substrings, each of which is delimited by the string position of its 5' end and the string position of its 3' end within the parent string. In a second dimension, each substring is associated with or pointed to one or multiple attribute fields. Such attribute fields contain annotations to the region on the nucleotide sequence represented by the substring.

For example, a sequence under investigation is 520 bases long and represented by a string named SeqTarget. There is a minor groove in the 5' upstream non-coding region from position 12 to 38, which is identified as a binding site for an enhancer protein HM-A, which in turn will increase the transcription of the gene represented by SeqTarget. Here, the substring is represented as (12, 38) and has the following attributes: [upstream uncoded], [minor groove], [HM-A binding] and [increase transcription upon binding by HM-A]. Similarly, other types of information can be stored and structured in this manner, such as information related to the whole sequence, *e.g.*, whether the sequence is a full length viral gene, a mammalian house keeping gene or an EST from clone X, information related to the 3' down stream non-coding region, *e.g.*, hair pin structure, and information related to various domains of the coding region, *e.g.*, Zinc finger.

This data structure is an open structure and is robust enough to accommodate newly generated data and acquired knowledge. Such a structure is also a flexible structure. It can be trimmed down to a 1-D string to facilitate data mining and analysis steps, such as

clustering, repeat-masking, and HMM analysis. Meanwhile, such a data structure also can extend the associated attributes into multiple dimensions. Pointers can be established among the dimensioned attributes when needed to facilitate data management and processing in a comprehensive genomics knowledgebase. Furthermore, such a data structure is object-oriented. Polymorphism can be represented by a family or class of sequence objects, each of which has an internal structure as discussed above. The common traits are abstracted and assigned to the parent object, whereas each child object represents a specific variant of the family or class. Such a data structure allows data to be efficiently retrieved, updated and integrated by the software applications associated with the sequence database and/or knowledgebase.

The present invention contemplates providing processes for effecting analysis and modeling, which are described in the following section.

Optionally, the present invention further contemplates that the machine of the present invention will embody in some manner a utility or function for the polynucleotide or polypeptide of the present invention. The function or utility of the polynucleotide or polypeptide can be a function or utility for the sequence data, *per se*, or of the tangible material. Exemplary function or utilities include the name (per International Union of Biochemistry and Molecular Biology rules of nomenclature) or function of the enzyme or protein represented by the polynucleotide or polypeptide of the present invention; the metabolic pathway of the protein represented by the polynucleotide or polypeptide of the present invention; the substrate or product or structural role of the protein represented by the polynucleotide or polypeptide of the present invention; or, the phenotype (e.g., an agronomic or pharmacological trait) affected by modulating expression or activity of the protein represented by the polynucleotide or polypeptide of the present invention.

B. Computer Analysis and Modeling

The present invention provides a process of modeling and analyzing data representative of a polynucleotide or polypeptide sequence of the present invention. The process comprises entering sequence data of a polynucleotide or polypeptide of the present invention into a machine having a hardware or software sequence modeling and analysis system, developing data structures to facilitate access to the sequence data, manipulating the data to model or analyze the structure or activity of the polynucleotide or polypeptide, and displaying the results of the modeling or analysis. Thus, the present invention provides a process for effecting the use, analysis, or modeling of a polynucleotide

sequence or its derived peptide sequence through use of a computer having a memory. The process comprises 1) placing into the memory data representing a polynucleotide having the sequence of a polynucleotide of the present invention,
developing within the memory a data structure associated with the data and reflecting the underlying organization and structure of the data to facilitate program access to data elements corresponding to logical sub-components of the sequence, 2) programming the computer with a program containing instructions sufficient to implement the process for effecting the use, analysis, or modeling of the polynucleotide sequence or the peptide sequence, and, 3) executing the program on the computer while granting the program access to the data and to the data structure within the memory.

A variety of modeling and analytic tools are well known in the art and available commercially. Included amongst the modeling/analysis tools are methods to: 1) recognize overlapping sequences (e.g., from a sequencing project) with a polynucleotide of the present invention and create an alignment called a "contig"; 2) identify restriction enzyme sites of a polynucleotide of the present invention; 3) identify the products of a T1 ribonuclease digestion of a polynucleotide of the present invention; 4) identify PCR primers with minimal self-complementarity; 5) compute pairwise distances between sequences in an alignment, reconstruct phylogentic trees using distance methods, and calculate the degree of divergence of two protein coding regions; 6) identify patterns such as coding regions, terminators, repeats, and other consensus patterns in polynucleotides of the present invention; 7) identify RNA secondary structure; 8) identify sequence motifs, isoelectric point, secondary structure, hydrophobicity, and antigenicity in polypeptides of the present invention; 9) translate polynucleotides of the present invention and backtranslate polypeptides of the present invention; and 10) compare two protein or nucleic acid sequences and identifying points of similarity or dissimilarity between them.

The processes for effecting analysis and modeling can be produced independently or obtained from commercial suppliers. Exemplary analysis and modeling tools are provided in products such as InforMax's (Bethesda, MD) Vector NTI Suite (Version 5.5), Intelligenetics' (Mountain View, CA) PC/Gene program, and Genetics Computer Group's (Madison, WI) Wisconsin Package (Version 10.0); these tools, and the functions they perform, (as provided and disclosed by the programs and accompanying literature) are incorporated herein by reference and are described in more detail in section C which follows.

Thus, in a further embodiment, the present invention provides a machine-readable medium containing a computer program and data, comprising a program stored on the medium containing instructions sufficient to implement a process for effecting the use, analysis, or modeling of a representation of a polynucleotide or peptide sequence. The data stored on the medium represents a sequence of a polynucleotide having the sequence of a polynucleotide of the present invention. The medium also includes a data structure reflecting the underlying organization and structure of the data to facilitate program access to data elements corresponding to logical sub-components of the sequence, the data structure being inherent in the program and in the way in which the program organizes and accesses the data.

C. Homology Searches

As an example of such a comparative analysis, the present invention provides a process of identifying a candidate homologue (i.e., an ortholog or paralog) of a polynucleotide or polypeptide of the present invention. The process comprises entering sequence data of a polynucleotide or polypeptide of the present invention into a machine having a hardware or software sequence analysis system, developing data structures to facilitate access to the sequence data, manipulating the data to analyze the structure the polynucleotide or polypeptide, and displaying the results of the analysis. A candidate homologue has statistically significant probability of having the same biological function (e.g., catalyzes the same reaction, binds to homologous proteins/nucleic acids, has a similar structural role) as the reference sequence to which it is compared. Accordingly, the polynucleotides and polypeptides of the present invention have utility in identifying homologs in animals or other plant species, particularly those in the family *Gramineae* such as, but not limited to, sorghum, wheat, or rice.

The process of the present invention comprises obtaining data representing a polynucleotide or polypeptide test sequence. Test sequences can be obtained from a nucleic acid of an animal or plant. Test sequences can be obtained directly or indirectly from sequence databases including, but not limited to, those such as: GenBank, EMBL, GenSeq, SWISS-PROT, or those available on-line via the UK Human Genome Mapping Project (HGMP) GenomeWeb. In some embodiments the test sequence is obtained from a plant species other than maize whose function is uncertain but will be compared to the test sequence to determine sequence similarity or sequence identity. The test sequence data is entered into a machine, such as a computer, containing: i) data representing a reference

sequence and, ii) a hardware or software sequence comparison system to compare the reference and test sequence for sequence similarity or identity.

Exemplary sequence comparison systems are provided for in sequence analysis software such as those provided by the Genetics Computer Group (Madison, WI) or InforMax (Bethesda, MD), or Intelligenetics (Mountain View, CA). Optionally, sequence comparison is established using the BLAST or GAP suite of programs. Generally, a smallest sum probability value (P(N)) of less than 0.1, or alternatively, less than 0.01, 0.001, 0.0001, or 0.00001 using the BLAST 2.0 suite of algorithms under default parameters identifies the test sequence as a candidate homologue (i.e., an allele, ortholog, or paralog) of the reference sequence. Those of skill in the art will recognize that a candidate homologue has an increased statistical probability of having the same or similar function as the gene/protein represented by the test sequence.

The reference sequence can be the sequence of a polypeptide or a polynucleotide of the present invention. The reference or test sequence is each optionally at least 25 amino acids or at least 100 nucleotides in length. The length of the reference or test sequences can be the length of the polynucleotide or polypeptide described, respectively, above in the sections entitled "Nucleic Acids" (particularly section (g)), and "Proteins". As those of skill in the art are aware, the greater the sequence identity/similarity between a reference sequence of known function and a test sequence, the greater the probability that the test sequence will have the same or similar function as the reference sequence. The results of the comparison between the test and reference sequences are outputted (e.g., displayed, printed, recorded) via any one of a number of output devices and/or media (e.g., computer monitor, hard copy, or computer readable medium).

Detection of Nucleic Acids

The present invention further provides methods for detecting a polynucleotide of the present invention in a nucleic acid sample suspected of containing a polynucleotide of the present invention, such as a plant cell lysate, particularly a lysate of maize. In some embodiments, a cognate gene of a polynucleotide of the present invention or portion thereof can be amplified prior to the step of contacting the nucleic acid sample with a polynucleotide of the present invention. The nucleic acid sample is contacted with the polynucleotide to form a hybridization complex. The polynucleotide hybridizes under stringent conditions to a gene encoding a polypeptide of the present invention. Formation of the hybridization complex is used to detect a gene encoding a polypeptide of the present

invention in the nucleic acid sample. Those of skill will appreciate that an isolated nucleic acid comprising a polynucleotide of the present invention should lack cross-hybridizing sequences in common with non-target genes that would yield a false positive result. Detection of the hybridization complex can be achieved using any number of well known methods. For example, the nucleic acid sample, or a portion thereof, may be assayed by hybridization formats including but not limited to, solution phase, solid phase, mixed phase, or *in situ* hybridization assays.

Detectable labels suitable for use in the present invention include any composition detectable by spectroscopic, radioisotopic, photochemical, biochemical, immunochemical, electrical, optical or chemical means. Useful labels in the present invention include biotin for staining with labeled streptavidin conjugate, magnetic beads, fluorescent dyes, radiolabels, enzymes, and colorimetric labels. Other labels include ligands which bind to antibodies labeled with fluorophores, chemiluminescent agents, and enzymes. Labeling the nucleic acids of the present invention is readily achieved such as by the use of labeled PCR primers.

Although the present invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be obvious that certain changes and modifications may be practiced within the scope of the appended claims.

Example 1

This example describes the construction of a cDNA library.

Total RNA can be isolated from maize tissues with TRIzol Reagent (Life Technology Inc. Gaithersburg, MD) using a modification of the guanidine isothiocyanate/acid-phenol procedure described by Chomczynski and Sacchi (Chomczynski, P., and Sacchi, N. *Anal. Biochem.* 162, 156 (1987)). In brief, plant tissue samples are pulverized in liquid nitrogen before the addition of the TRIzol Reagent, and then further homogenized with a mortar and pestle. Addition of chloroform followed by centrifugation is conducted for separation of an aqueous phase and an organic phase. The total RNA is recovered by precipitation with isopropyl alcohol from the aqueous phase.

The selection of poly(A)+ RNA from total RNA can be performed using PolyATact system (Promega Corporation. Madison, WI). Biotinylated oligo(dT) primers are used to hybridize to the 3' poly(A) tails on mRNA. The hybrids are captured using streptavidin coupled to paramagnetic particles and a magnetic separation stand. The

mRNA is then washed at high stringency conditions and eluted by RNase-free deionized water.

cDNA synthesis and construction of unidirectional cDNA libraries can be accomplished using the SuperScript Plasmid System (Life Technology Inc. Gaithersburg, MD). The first strand of cDNA is synthesized by priming an oligo(dT) primer containing a Not I site. The reaction is catalyzed by SuperScript Reverse Transcriptase II at 45°C. The second strand of cDNA is labeled with alpha-^{32}P-dCTP and a portion of the reaction analyzed by agarose gel electrophoresis to determine cDNA sizes. cDNA molecules smaller than 500 base pairs and unligated adapters are removed by Sephacryl-S400 chromatography. The selected cDNA molecules are ligated into pSPORT1 vector in between of *Not* I and *Sal* I sites.

Alternatively, cDNA libraries can be prepared by any one of many methods available. For example, the cDNAs may be introduced into plasmid vectors by first preparing the cDNA libraries in Uni-ZAP™ XR vectors according to the manufacturer's protocol (Stratagene Cloning Systems, La Jolla, CA). The Uni-ZAP™ XR libraries are converted into plasmid libraries according to the protocol provided by Stratagene. Upon conversion, cDNA inserts will be contained in the plasmid vector pBluescript. In addition, the cDNAs may be introduced directly into precut Bluescript II SK(+) vectors (Stratagene) using T4 DNA ligase (New England Biolabs), followed by transfection into DH10B cells according to the manufacturer's protocol (GIBCO BRL Products). Once the cDNA inserts are in plasmid vectors, plasmid DNAs are prepared from randomly picked bacterial colonies containing recombinant pBluescript plasmids, or the insert cDNA sequences are amplified via polymerase chain reaction using primers specific for vector sequences flanking the inserted cDNA sequences. Amplified insert DNAs or plasmid DNAs are sequenced in dye-primer sequencing reactions to generate partial cDNA sequences (expressed sequence tags or "ESTs"; see Adams *et al.*, (1991) *Science 252*:1651-1656). The resulting ESTs are analyzed using a Perkin Elmer Model 377 fluorescent sequencer.

Example 2

This method describes construction of a full-length enriched cDNA library.

An enriched full-length cDNA library can be constructed using one of two variations of the method of Carninci *et al. Genomics* 37: 327-336, 1996. These variations are based on chemical introduction of a biotin group into the diol residue of the 5' cap structure of eukaryotic mRNA to select full-length first strand cDNA. The selection

occurs by trapping the biotin residue at the cap sites using streptavidin-coated magnetic beads followed by RNase I treatment to eliminate incompletely synthesized cDNAs. Second strand cDNA is synthesized using established procedures such as those provided in Life Technologies' (Rockville, MD) "SuperScript Plasmid System for cDNA Synthesis and Plasmid Cloning" kit. Libraries made by this method have been shown to contain 50% to 70% full-length cDNAs.

The first strand synthesis methods are detailed below. An asterisk denotes that the reagent was obtained from Life Technologies, Inc.

A. First strand cDNA synthesis method 1 (with trehalose)

mRNA (10ug)	25µl
*Not I primer (5ug)	10µl
*5x 1st strand buffer	43µl
*0.1m DTT	20µl
*dNTP mix 10mm	10µl
BSA 10ug/µl	1µl
Trehalose (saturated)	59.2µl
RNase inhibitor (Promega)	1.8µl
*Superscript II RT 200u/µl	20µl
100 % glycerol	18µl
Water	7µl

The mRNA and Not I primer are mixed and denatured at 65°C for 10 min. They are then chilled on ice and other components added to the tube. Incubation is at 45°C for 2 min. Twenty microliters of RT (reverse transcriptase) is added to the reaction and start program on the thermocycler (MJ Research, Waltham, MA):

Step 1	45°C 10min
Step 2	45°C –0.3°C/cycle , 2 seconds/cycle
Step 3	go to 2 for 33 cycles
Step 4	35°C 5min
Step 5	45°C 5min
Step 6	45°C 0.2°C/cycle, 1 sec/cycle
Step 7	go to 7 for 49 cycles

Step 8	55°C 0.1°C/cycle, 12 sec/cycle	
Step 9	go to 8 for 49 cycles	
Step 10	55°C 2min	
Step11	60°C 2min	
Step 12	go to 11 for 9 times	
Step 13	4°C forever	
Step14	end	

B. First strand cDNA synthesis method 2

mRNA (10μg)	25μl
water	30μl
*Not I adapter primer (5μg)	10μl

65°C for 10min, chill on ice, then add following reagents,

*5x first buffer	20μl
*0.1M DTT	10μl
*10mM dNTP mix	5μl

Incubate at 45°C for 2min, then add 10μl of *Superscript II RT (200u/μl), start the following program:

Step 1	45°C for 6 sec, -0.1°C/cycle
Step 2	go to 1 for 99 additional cycles
Step 3	35°C for 5min
Step 4	45°C for 60 min
Step 5	50°C for 10 min
Step 6	4°C forever
Step 7	end

After the 1st strand cDNA synthesis, the DNA is extracted by phenol according to standard procedures, and then precipitated in NaOAc and ethanol, and stored in –20°C.

C. Oxidization of the diol group of mRNA for biotin labeling

First strand cDNA is spun down and washed once with 70% EtOH. The pellet resuspended in 23.2 μl of DEPC treated water and put on ice. Prepare 100 mM of NaIO4 freshly, and then add the following reagents:

mRNA:1st cDNA (start with 20µg mRNA)	46.4µl
100mM NaIO4 (freshly made)	2.5µl
NaOAc 3M pH4.5	1.1µl

To make 100 mM NaIO4, use 21.39µg of NaIO4 for 1µl of water.

Wrap the tube in a foil and incubate on ice for 45min.

After the incubation, the reaction is then precipitated in:

5M NaCl	10µl
20%SDS	0.5µl
isopropanol	61µl

Incubate on ice for at least 30 min, then spin it down at max speed at 4°C for 30 min and wash once with 70% ethanol and then 80% EtOH.

D. Biotinylation of the mRNA diol group

Resuspend the DNA in 110µl DEPC treated water, then add the following reagents:

20% SDS	5 µl
2 M NaOAc pH 6.1	5 µl
10mm biotin hydrazide (freshly made)	300 µl

Wrap in a foil and incubate at room temperature overnight.

E. RNase I treatment

Precipitate DNA in:

5M NaCl	10µl
2M NaOAc pH 6.1	75µl
biotinylated mRNA:cDNA	420µl
100% EtOH (2.5Vol)	1262.5µl

(Perform this precipitation in two tubes and split the 420 µl of DNA into 210 µl each, add 5µl of 5M NaCl, 37.5µl of 2M NaOAc pH 6.1, and 631.25 µl of 100% EtOH).

Store at –20°C for at least 30 min. Spin the DNA down at 4°C at maximal speed for 30 min. and wash with 80% EtOH twice, then dissolve DNA in 70µl RNase free water. Pool two tubes and end up with 140 µl.

Add the following reagents:

RNase One 10U/μl	40μl
1st cDNA:RNA	140μl
10X buffer	20μl

Incubate at 37°C for 15min.

Add 5μl of 40μg/μl yeast tRNA to each sample for capturing.

F. Full length 1st cDNA capturing

Blocking the beads with yeast tRNA:

Beads	1ml
Yeast tRNA 40μg/μl	5μl

Incubate on ice for 30min with mixing, wash 3 times with 1ml of 2M NaCl , 50mmEDTA, pH 8.0.

Resuspend the beads in 800μl of 2M NaCl , 50mm EDTA, pH 8.0, add RNase I treated sample 200μl, and incubate the reaction for 30min at room temperature.

Capture the beads using the magnetic stand, save the supernatant, and start following washes:

2 washes with 2M NaCl , 50mm EDTA, pH 8.0, 1 ml each time,

1 wash with 0.4% SDS, 50μg/ml tRNA,

1 wash with 10mm Tris-Cl pH 7.5, 0.2mm EDTA, 10mm NaCl, 20% glycerol,

1 wash with 50μg/ml tRNA,

1 wash with 1st cDNA buffer

G. Second strand cDNA synthesis

Resuspend the beads in:

*5X first buffer	8μl
*0.1mM DTT	4μl
*10mm dNTP mix	8μl
*5X 2nd buffer	60μl
*E.coli Ligase 10U/μl	2μl
*E.coli DNA polymerase 10U/μl	8μl
*E. coli RNaseH 2U/μl	2μl
P32 dCTP 10μci/μl	2μl
Or water up to 300μl	208μl

Incubate at 16°C for 2hr with mixing the reaction in every 30 min.

Add 4µl of T4 DNA polymerase and incubate for additional 5 min at 16°C.

Elute 2nd cDNA from the beads.

Use a magnetic stand to separate the 2nd cDNA from the beads, then resuspend the beads in 200µl of water, and then separate again, pool the samples (about 500µl).

Add 200 µl of water to the beads, then 200µl of phenol:chloroform, vortex, and spin to separate the sample with phenol.

Pool the DNA together (about 700µl) and use phenol to clean the DNA again, DNA is then precipitated in 2µg of glycogen and 0.5 vol of 7.5M NH4OAc and 2 vol of 100% EtOH. Precipitate overnight. Spin down the pellet and wash with 70% EtOH, air-dry the pellet.

	DNA	250µl	DNA	200µl
	7.5M NH4OAc	125µl	7.5M NH4OAc	100µl
	100% EtOH	750µl	100% EtOH	600µl
	glycogen 1µg/µl	2µl	glycogen 1µg/µl	2µl

H. Sal I adapter ligation

Resuspend the pellet in 26 µl of water and use 1µl for TAE gel.

Set up reaction as following:

2nd strand cDNA	25µl
*5X T4 DNA ligase buffer	10µl
*Sal I adapters	10µl
*T4 DNA ligase	5µl

Mix gently, incubate the reaction at 16°C overnight.

Add 2µl of ligase second day and incubate at room temperature for 2 hrs (optional).

Add 50µl water to the reaction and use 100µl of phenol to clean the DNA, 90µl of the upper phase is transferred into a new tube and precipitate in:

Glycogen 1µg/µl	2µl
Upper phase DNA	90µl

7.5M NH4OAc 50µl

100% EtOH 300µl

precipitate at –20°C overnight

Spin down the pellet at 4°C and wash in 70% EtOH, dry the pellet.

I. Not I digestion

2nd cDNA 41µl

*Reaction 3 buffer 5µl

*Not I 15u/µl 4µl

Mix gently and incubate the reaction at 37°C for 2hr.

Add 50 µl of water and 100µl of phenol, vortex , and take 90µl of the upper phase to a new tube, then add 50µl of NH40Ac and 300 µl of EtOH. Precipitate overnight at -20°C.

Cloning, ligation, and transformation are performed per the Superscript cDNA synthesis kit.

Example 3

This example describes cDNA sequencing and library subtraction.

Individual colonies can be picked and DNA prepared either by PCR with M13 forward primers and M13 reverse primers, or by plasmid isolation. cDNA clones can be sequenced using M13 reverse primers.

cDNA libraries are plated out on 22 x 22 cm^2 agar plate at density of about 3,000 colonies per plate. The plates are incubated in a 37°C incubator for 12-24 hours. Colonies are picked into 384-well plates by a robot colony picker, Q-bot (GENETIX Limited). These plates are incubated overnight at 37°C. Once sufficient colonies are picked, they are pinned onto 22 x 22 cm^2 nylon membranes using Q-bot. Each membrane holds 9,216 or 36,864 colonies. These membranes are placed onto an agar plate with an appropriate antibiotic. The plates are incubated at 37°C overnight.

After colonies are recovered on the second day, these filters are placed on filter paper prewetted with denaturing solution for four minutes, then incubated on top of a boiling water bath for an additional four minutes. The filters are then placed on filter paper prewetted with neutralizing solution for four minutes. After excess solution is removed by placing the filters on dry filter papers for one minute, the colony side of the

filters is placed into Proteinase K solution, incubated at 37°C for 40-50 minutes. The filters are placed on dry filter papers to dry overnight. DNA is then cross-linked to nylon membrane by UV light treatment.

Colony hybridization is conducted as described by Sambrook,J., Fritsch, E.F. and Maniatis, T., (in Molecular Cloning: A laboratory Manual, 2nd Edition). The following probes can be used in colony hybridization:

1. First strand cDNA from the same tissue as the library was made from to remove the most redundant clones.

2. 48-192 most redundant cDNA clones from the same library based on previous sequencing data.

3. 192 most redundant cDNA clones in the entire maize sequence database.

4. A Sal-A20 oligo nucleotide: TCG ACC CAC GCG TCC GAA AAA AAA AAA AAA AAA AAA AAA, removes clones containing a poly A tail but no cDNA.

5. cDNA clones derived from rRNA.

The image of the autoradiography is scanned into computer and the signal intensity and cold colony addresses of each colony is analyzed. Re-arraying of cold-colonies from 384 well plates to 96 well plates is conducted using Q-bot.

Example 4

This example describes identification of the gene from a computer homology search.

Gene identities can be determined by conducting BLAST (Basic Local Alignment Search Tool; Altschul, S. F., et al., (1993) J. Mol. Biol. 215:403-410; see also www.ncbi.nlm.nih.gov/BLAST/) searches under default parameters for similarity to sequences contained in the BLAST "nr" database (comprising all non-redundant GenBank CDS translations, sequences derived from the 3-dimensional structure Brookhaven Protein Data Bank, the last major release of the SWISS-PROT protein sequence database, EMBL, and DDBJ databases). The cDNA sequences are analyzed for similarity to all publicly available DNA sequences contained in the "nr" database using the BLASTN algorithm. The DNA sequences are translated in all reading frames and compared for similarity to all publicly available protein sequences contained in the "nr" database using the BLASTX algorithm (Gish, W. and States, D. J. *Nature Genetics 3*:266-272 (1993)) provided by the NCBI. In some cases, the sequencing data from two or more clones containing overlapping segments of DNA are used to construct contiguous DNA sequences.

Sequence alignments and percent identity calculations can be performed using the Megalign program of the LASERGENE bioinformatics computing suite (DNASTAR Inc., Madison, WI). Multiple alignment of the sequences can be performed using the Clustal method of alignment (Higgins and Sharp (1989) *CABIOS.* 5:151-153) with the default parameters (GAP PENALTY=10, GAP LENGTH PENALTY=10). Default parameters for pairwise alignments using the Clustal method are KTUPLE 1, GAP PENALTY=3, WINDOW=5 and DIAGONALS SAVED=5.

Example 5

This example describes expression of transgenes in monocot cells.

A transgene comprising a cDNA encoding the instant polypeptides in sense orientation with respect to the maize 27 kD zein promoter that is located 5' to the cDNA fragment, and the 10 kD zein 3' end that is located 3' to the cDNA fragment, can be constructed. The cDNA fragment of this gene may be generated by polymerase chain reaction (PCR) of the cDNA clone using appropriate oligonucleotide primers. Cloning sites (NcoI or SmaI) can be incorporated into the oligonucleotides to provide proper orientation of the DNA fragment when inserted into the digested vector pML103 as described below. Amplification is then performed in a standard PCR. The amplified DNA is then digested with restriction enzymes NcoI and SmaI and fractionated on an agarose gel. The appropriate band can be isolated from the gel and combined with a 4.9 kb NcoI-SmaI fragment of the plasmid pML103. Plasmid pML103 has been deposited under the terms of the Budapest Treaty at ATCC (American Type Culture Collection, 10801 University Blvd., Manassas, VA 20110-2209), and bears accession number ATCC 97366. The DNA segment from pML103 contains a 1.05 kb SalI-NcoI promoter fragment of the maize 27 kD zein gene and a 0.96 kb SmaI-SalI fragment from the 3' end of the maize 10 kD zein gene in the vector pGem9Zf(+) (Promega). Vector and insert DNA can be ligated at 15°C overnight, essentially as described (Maniatis). The ligated DNA may then be used to transform *E. coli* XL1-Blue (Epicurian Coli XL-1 Blue; Stratagene). Bacterial transformants can be screened by restriction enzyme digestion of plasmid DNA and limited nucleotide sequence analysis using the dideoxy chain termination method (Sequenase DNA Sequencing Kit; U. S. Biochemical). The resulting plasmid construct would comprise a transgene encoding, in the 5' to 3' direction, the maize 27 kD zein promoter, a cDNA fragment encoding the instant polypeptides, and the 10 kD zein 3' region.

The transgene described above can then be introduced into maize cells by the following procedure. Immature maize embryos can be dissected from developing caryopses derived from crosses of the inbred maize lines H99 and LH132. The embryos are isolated 10 to 11 days after pollination when they are 1.0 to 1.5 mm long. The embryos are then placed with the axis-side facing down and in contact with agarose-solidified N6 medium (Chu *et al.* (1975) *Sci. Sin. Peking* 18:659-668). The embryos are kept in the dark at 27°C. Friable embryogenic callus consisting of undifferentiated masses of cells with somatic proembryoids and embryoids borne on suspensor structures proliferates from the scutellum of these immature embryos. The embryogenic callus isolated from the primary explant can be cultured on N6 medium and sub-cultured on this medium every 2 to 3 weeks.

The plasmid, p35S/Ac (Hoechst Ag, Frankfurt, Germany) or equivalent may be used in transformation experiments in order to provide for a selectable marker. This plasmid contains the *Pat* gene (see European Patent Publication 0 242 236) which encodes phosphinothricin acetyl transferase (PAT). The enzyme PAT confers resistance to herbicidal glutamine synthetase inhibitors such as phosphinothricin. The *pat* gene in p35S/Ac is under the control of the 35S promoter from Cauliflower Mosaic Virus (Odell et al. (1985) *Nature* 313:810-812) and the 3' region of the nopaline synthase gene from the T-DNA of the Ti plasmid of *Agrobacterium tumefaciens*.

The particle bombardment method (Klein *et al.* (1987) *Nature* 327:70-73) may be used to transfer genes to the callus culture cells. According to this method, gold particles (1 μm in diameter) are coated with DNA using the following technique. Ten μg of plasmid DNAs are added to 50 μL of a suspension of gold particles (60 mg per mL). Calcium chloride (50 μL of a 2.5 M solution) and spermidine free base (20 μL of a 1.0 M solution) are added to the particles. The suspension is vortexed during the addition of these solutions. After 10 minutes, the tubes are briefly centrifuged (5 sec at 15,000 rpm) and the supernatant removed. The particles are resuspended in 200 μL of absolute ethanol, centrifuged again and the supernatant removed. The ethanol rinse is performed again and the particles resuspended in a final volume of 30 μL of ethanol. An aliquot (5 μL) of the DNA-coated gold particles can be placed in the center of a Kapton flying disc (Bio-Rad Labs). The particles are then accelerated into the maize tissue with a Biolistic PDS-1000/He (Bio-Rad Instruments, Hercules CA), using a helium pressure of 1000 psi, a gap distance of 0.5 cm and a flying distance of 1.0 cm.

For bombardment, the embryogenic tissue is placed on filter paper over agarose-solidified N6 medium. The tissue is arranged as a thin lawn and covers a circular area of about 5 cm in diameter. The petri dish containing the tissue can be placed in the chamber of the PDS-1000/He approximately 8 cm from the stopping screen. The air in the chamber is then evacuated to a vacuum of 28 inches of Hg. The macrocarrier is accelerated with a helium shock wave using a rupture membrane that bursts when the He pressure in the shock tube reaches 1000 psi.

Seven days after bombardment the tissue can be transferred to N6 medium that contains gluphosinate (2 mg per liter) and lacks casein or proline. The tissue continues to grow slowly on this medium. After an additional 2 weeks the tissue can be transferred to fresh N6 medium containing gluphosinate. After 6 weeks, areas of about 1 cm in diameter of actively growing callus can be identified on some of the plates containing the glufosinate-supplemented medium. These calli may continue to grow when sub-cultured on the selective medium.

Plants can be regenerated from the transgenic callus by first transferring clusters of tissue to N6 medium supplemented with 0.2 mg per liter of 2,4-D. After two weeks the tissue can be transferred to regeneration medium (Fromm *et al.* (1990) *Bio/Technology 8*:833-839).

Example 6

This example describes expression of transgenes in dicot cells.

A seed-specific expression cassette composed of the promoter and transcription terminator from the gene encoding the β subunit of the seed storage protein phaseolin from the bean *Phaseolus vulgaris* (Doyle *et al.* (1986) *J. Biol. Chem. 261*:9228-9238) can be used for expression of the instant polypeptides in transformed soybean. The phaseolin cassette includes about 500 nucleotides upstream (5') from the translation initiation codon and about 1650 nucleotides downstream (3') from the translation stop codon of phaseolin. Between the 5' and 3' regions are the unique restriction endonuclease sites Nco I (which includes the ATG translation initiation codon), SmaI, KpnI and XbaI. The entire cassette is flanked by Hind III sites.

The cDNA fragment of this gene may be generated by polymerase chain reaction (PCR) of the cDNA clone using appropriate oligonucleotide primers. Cloning sites can be incorporated into the oligonucleotides to provide proper orientation of the DNA fragment when inserted into the expression vector. Amplification is then performed as described

above, and the isolated fragment is inserted into a pUC18 vector carrying the seed expression cassette.

Soybean embroys may then be transformed with the expression vector comprising sequences encoding the instant polypeptides. To induce somatic embryos, cotyledons, 3-5 mm in length dissected from surface sterilized, immature seeds of the soybean cultivar A2872, can be cultured in the light or dark at 26°C on an appropriate agar medium for 6-10 weeks. Somatic embryos which produce secondary embryos are then excised and placed into a suitable liquid medium. After repeated selection for clusters of somatic embryos which multiplied as early, globular staged embryos, the suspensions are maintained as described below.

Soybean embryogenic suspension cultures can be maintained in 35 mL liquid media on a rotary shaker, 150 rpm, at 26°C with florescent lights on a 16:8 hour day/night schedule. Cultures are subcultured every two weeks by inoculating approximately 35 mg of tissue into 35 mL of liquid medium.

Soybean embryogenic suspension cultures may then be transformed by the method of particle gun bombardment (Klein *et al.* (1987) *Nature* (London) *327*:70-73, U.S. Patent No. 4,945,050). A Du Pont Biolistic PDS1000/HE instrument (helium retrofit) can be used for these transformations.

A selectable marker gene which can be used to facilitate soybean transformation is a transgene composed of the 35S promoter from Cauliflower Mosaic Virus (Odell *et al.*(1985) *Nature 313*:810-812), the hygromycin phosphotransferase gene from plasmid pJR225 (from *E. coli*; Gritz *et al.*(1983) *Gene 25*:179-188) and the 3' region of the nopaline synthase gene from the T-DNA of the Ti plasmid of *Agrobacterium tumefaciens*. The seed expression cassette comprising the phaseolin 5' region, the fragment encoding the instant polypeptides and the phaseolin 3' region can be isolated as a restriction fragment. This fragment can then be inserted into a unique restriction site of the vector carrying the marker gene.

To 50 µL of a 60 mg/mL 1 µm gold particle suspension is added (in order): 5 µL DNA (1 µg/µL), 20 µl spermidine (0.1 M), and 50 µL $CaCl_2$ (2.5 M). The particle preparation is then agitated for three minutes, spun in a microfuge for 10 seconds and the supernatant removed. The DNA-coated particles are then washed once in 400 µL 70% ethanol and resuspended in 40 µL of anhydrous ethanol. The DNA/particle suspension can be sonicated three times for one second each. Five microliters of the DNA-coated gold particles are then loaded on each macro carrier disk.

Approximately 300-400 mg of a two-week-old suspension culture is placed in an empty 60x15 mm petri dish and the residual liquid removed from the tissue with a pipette. For each transformation experiment, approximately 5-10 plates of tissue are normally bombarded. Membrane rupture pressure is set at 1100 psi and the chamber is evacuated to a vacuum of 28 inches mercury. The tissue is placed approximately 3.5 inches away from the retaining screen and bombarded three times. Following bombardment, the tissue can be divided in half and placed back into liquid and cultured as described above.

Five to seven days post bombardment, the liquid media may be exchanged with fresh media, and eleven to twelve days post bombardment with fresh media containing 50 mg/mL hygromycin. This selective media can be refreshed weekly. Seven to eight weeks post bombardment, green, transformed tissue may be observed growing from untransformed, necrotic embryogenic clusters. Isolated green tissue is removed and inoculated into individual flasks to generate new, clonally propagated, transformed embryogenic suspension cultures. Each new line may be treated as an independent transformation event. These suspensions can then be subcultured and maintained as clusters of immature embryos or regenerated into whole plants by maturation and germination of individual somatic embryos.

Example 7

This example describes expression of a transgene in microbial cells.

The cDNAs encoding the instant polypeptides can be inserted into the T7 *E. coli* expression vector pBT430. This vector is a derivative of pET-3a (Rosenberg *et al.* (1987) *Gene 56*:125-135) which employs the bacteriophage T7 RNA polymerase/T7 promoter system. Plasmid pBT430 was constructed by first destroying the EcoR I and Hind III sites in pET-3a at their original positions. An oligonucleotide adaptor containing EcoR I and Hind III sites was inserted at the BamH I site of pET-3a. This created pET-3aM with additional unique cloning sites for insertion of genes into the expression vector. Then, the Nde I site at the position of translation initiation was converted to an Nco I site using oligonucleotide-directed mutagenesis. The DNA sequence of pET-3aM in this region, 5'-CATATGG, was converted to 5'-CCCATGG in pBT430.

Plasmid DNA containing a cDNA may be appropriately digested to release a nucleic acid fragment encoding the protein. This fragment may then be purified on a 1% NuSieve GTG low melting agarose gel (FMC). Buffer and agarose contain 10 µg/ml ethidium bromide for visualization of the DNA fragment. The fragment can then be purified from

the agarose gel by digestion with GELase (Epicentre Technologies) according to the manufacturer's instructions, ethanol precipitated, dried and resuspended in 20 μL of water. Appropriate oligonucleotide adapters may be ligated to the fragment using T4 DNA ligase (New England Biolabs, Beverly, MA). The fragment containing the ligated adapters can be purified from the excess adapters using low melting agarose as described above. The vector pBT430 is digested, dephosphorylated with alkaline phosphatase (NEB) and deproteinized with phenol/chloroform as described above. The prepared vector pBT430 and fragment can then be ligated at 16°C for 15 hours followed by transformation into DH5 electrocompetent cells (GIBCO BRL). Transformants can be selected on agar plates containing LB media and 100 μg/mL ampicillin. Transformants containing the gene encoding the instant polypeptides are then screened for the correct orientation with respect to the T7 promoter by restriction enzyme analysis.

For high level expression, a plasmid clone with the cDNA insert in the correct orientation relative to the T7 promoter can be transformed into *E. coli* strain BL21(DE3) (Studier et al. (1986) *J. Mol. Biol. 189*:113-130). Cultures are grown in LB medium containing ampicillin (100 mg/L) at 25°C. At an optical density at 600 nm of approximately 1, IPTG (isopropylthio-β-galactoside, the inducer) can be added to a final concentration of 0.4 mM and incubation can be continued for 3 h at 25°. Cells are then harvested by centrifugation and re-suspended in 50 μL of 50 mM Tris-HCl at pH 8.0 containing 0.1 mM DTT and 0.2 mM phenyl methylsulfonyl fluoride. A small amount of 1 mm glass beads can be added and the mixture sonicated 3 times for about 5 seconds each time with a microprobe sonicator. The mixture is centrifuged and the protein concentration of the supernatant determined. One microgram of protein from the soluble fraction of the culture can be separated by SDS-polyacrylamide gel electrophoresis. Gels can be observed for protein bands migrating at the expected molecular weight.

The above examples are provided to illustrate the invention but not to limit its scope. Other variants of the invention will be readily apparent to one of ordinary skill in the art and are encompassed by the appended claims. All publications, patents, patent applications, and computer programs cited herein are hereby incorporated by reference.

Deposits

Plasmids containing polynucleotide sequences of the invention were deposited on March 21, 2000, with the American Type Culture Collection (ATCC), 10801 University

Boulevard, Manassas, Virginia USA, 20110-2209, and assigned Accession Nos. PTA-1539, PTA-1540, and PTA-1541. These deposits will be maintained under the terms of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure. In addition, during the pendency of this patent application, access to the deposited cultures will be available to the Commissioner of Patents and Trademarks and to persons determined by the Commissioner to be entitled thereto under 37 C.F.R. §114 and 35 U.S.C. §122.

These deposits were made merely as a convenience for those of skill in the art and are not an admission that a deposit is required under 35 U.S.C. §112. All restrictions imposed by the depositor on the availability to the public of the deposited material will be irrevocably removed upon granting of a patent. However, it should be understood that the availability of a deposit does not constitute a license to practice the subject invention in derogation of patent rights granted by government action.

WHAT IS CLAIMED IS:

1. An isolated nucleic acid comprising a member selected from the group consisting of:

(a) a polynucleotide having at least 80% sequence identity, as determined by the GAP algorithm under default parameters, to a polynucleotide of SEQ ID NO: 1, 5, 9, 13, 17, 21, 25, 29, 33, 37, 41, 45, 49, 53, 57, 61, or 65;

(b) a polynucleotide encoding a polypeptide of SEQ ID NO: 2, 6, 10, 14, 18, 22, 26, 30, 34, 38, 42, 46, 50, 54, 58, 62, or 66;

(c) a polynucleotide amplified from a *Zea mays* nucleic acid library using primers which selectively hybridize, under stringent hybridization conditions, to loci within a polynucleotide of SEQ ID NO: 1, 5, 9, 13, 17, 21, 25, 29, 33, 37, 41, 45, 49, 53, 57, 61, or 65;

(d) a polynucleotide which selectively hybridizes, under stringent hybridization conditions and a wash in 0.1X SSC at 60°C, to a polynucleotide of SEQ ID NO: 1, 5, 9, 13, 17, 21, 25, 29, 33, 37, 41, 45, 49, 53, 57, 61, or 65;

(e) a polynucleotide of SEQ ID NO: 1, 5, 9, 13, 17, 21, 25, 29, 33, 37, 41, 45, 49, 53, 57, 61, or 65;

(f) a polynucleotide which is complementary to a polynucleotide of (a), (b), (c), (d), or (e); and

(g) a polynucleotide comprising at least 25 contiguous nucleotides from a polynucleotide of (a), (b), (c), (d), (e), or (f).

2. A recombinant expression cassette, comprising a member of claim 1 operably linked, in sense or anti-sense orientation, to a promoter.

3. A host cell comprising the recombinant expression cassette of claim 2.

4. A transgenic plant comprising a recombinant expression cassette of claim 2.

5. The transgenic plant of claim 4, wherein said plant is a monocot.

6. The transgenic plant of claim 4, wherein said plant is a dicot.

7. The transgenic plant of claim 4, wherein said plant is selected from the group consisting of: maize, soybean, sunflower, sorghum, canola, wheat, alfalfa, cotton, rice, barley, millet, peanut, and cocoa.

8. A transgenic seed from the transgenic plant of claim 4.

9. A method of modulating the level of signal transduction gene expression in a plant cell, comprising:

(a) introducing into a plant cell a recombinant expression cassette comprising a polynucleotide of claim 1 operably linked to a promoter;

(b) culturing the plant cell under plant cell growing conditions; and

(c) inducing expression of said polynucleotide for a time sufficient to modulate the level of signal transduction gene expression in said plant cell.

10. The method of claim 9, wherein the plant cell is from maize, wheat, rice, or soybean.

11. A method of modulating the level of signal transduction gene expression in a plant, comprising:

(a) introducing into a plant cell a recombinant expression cassette comprising a polynucleotide of claim 1 operably linked to a promoter;

(b) culturing the plant cell under plant cell growing conditions;

(c) regenerating a plant from said plant cell; and

(d) inducing expression of said polynucleotide for a time sufficient to modulate the level of signal transduction gene expression in said plant.

12. The method of claim 11, wherein the plant is maize, wheat, rice, or soybean.

13. An isolated protein comprising a member selected from the group consisting of:

(a) a polypeptide of at least 20 contiguous amino acids from a polypeptide of SEQ ID NO: 2, 6, 10, 14, 18, 22, 26, 30, 34, 38, 42, 46, 50, 54, 58, 62, or 66;

(b) a polypeptide of SEQ ID NO: 2, 6, 10, 14, 18, 22, 26, 30, 34, 38, 42, 46, 50, 54, 58, 62, or 66;

(c) a polypeptide having at least 80% sequence identity to, and having at least one epitope in common with, a polypeptide of SEQ ID NO: 2, 6, 10, 14, 18, 22, 26, 30, 34, 38, 42, 46, 50, 54, 58, 62, or 66, wherein said sequence identity is determined by the GAP algorithm under default parameters; and,

(d) at least one polypeptide encoded by a member of claim 1.

SEQUENCE LISTING

<110> Pioneer Hi-Bred International, Inc.

<120> Signal Transduction Genes and Methods of Use

<130> 1083-PCT

<150> US 60/134,292
<151> 1999-05-14

<150> US 60/142,996
<151> 1999-07-08

<160> 69

<170> FastSEQ for Windows Version 3.0

<210> 1
<211> 2029
<212> DNA
<213> Zea mays

<220>
<221> CDS
<222> (236)...(1543)

<400> 1

```
gtcgacccac gcgtccggaa gactgagtaa tgtagatttc tcatctgctg caatctctat     60
ccaagatacc aatgtatggt cagcacggtc tctctggtct tctgatgcga tgggatcttc    120
acctccttgt gattcacctg acaagtttgc aggtggtcag gagcgttctc ctctctctag    180
tccactgaga agtccagttt taaaatcaag aaacccaagt gcacctccat cacca atg     238
                                                            Met
                                                             1

cat cca aag ttg ttc ccg gag aac cat gtt tct cgc cct gat ggc aat      286
His Pro Lys Leu Phe Pro Glu Asn His Val Ser Arg Pro Asp Gly Asn
            5                  10                  15

gga agt gct agt ttt cac cca tta ccg ctc cct cct gct tct gta agc      334
Gly Ser Ala Ser Phe His Pro Leu Pro Leu Pro Pro Ala Ser Val Ser
        20                  25                  30

cca aag cag aca aat gct agc cac cag tta gtt cca aaa gct gag atg      382
Pro Lys Gln Thr Asn Ala Ser His Gln Leu Val Pro Lys Ala Glu Met
    35                  40                  45

cct tcg gtg gct ggt caa tgg cag aaa gga aaa ctc ttg ggt agt ggc      430
Pro Ser Val Ala Gly Gln Trp Gln Lys Gly Lys Leu Leu Gly Ser Gly
50                  55                  60                  65

acc ttt gga tgc gtg tat gag gcc acc aat agg cac act gga gct ctt      478
Thr Phe Gly Cys Val Tyr Glu Ala Thr Asn Arg His Thr Gly Ala Leu
                70                  75                  80

tgt gcc atg aaa gag gtg aat ata att cca gat gat gct aaa tca gtc      526
Cys Ala Met Lys Glu Val Asn Ile Ile Pro Asp Asp Ala Lys Ser Val
            85                  90                  95

gag tct ctg aag caa ttg gaa cag gaa att aag ttt ctt agt caa ttt      574
```

```
Glu Ser Leu Lys Gln Leu Glu Gln Glu Ile Lys Phe Leu Ser Gln Phe
        100                 105                 110

aag cat gaa aat att gtg cag tac tat ggc agt gaa act atc gaa gat      622
Lys His Glu Asn Ile Val Gln Tyr Tyr Gly Ser Glu Thr Ile Glu Asp
    115                 120                 125

aga ttc tac ata tac ctg gag tat gtt cat cct ggt tca atc cat aag      670
Arg Phe Tyr Ile Tyr Leu Glu Tyr Val His Pro Gly Ser Ile His Lys
130                 135                 140                 145

tat gtt cat caa cac tgt gga tcg ctg aca gaa gca gtc atc cgt aac      718
Tyr Val His Gln His Cys Gly Ser Leu Thr Glu Ala Val Ile Arg Asn
                150                 155                 160

ttc aca cgt cat att ctc aag ggt ctt gct ttt tta cat agc caa aag      766
Phe Thr Arg His Ile Leu Lys Gly Leu Ala Phe Leu His Ser Gln Lys
            165                 170                 175

att atg cat agg gat atc aaa ggg gca aat ttg tta gtt gat att aac      814
Ile Met His Arg Asp Ile Lys Gly Ala Asn Leu Leu Val Asp Ile Asn
        180                 185                 190

ggt gta gtg aaa ttg gct gac ttc gga atg gca aag cat ttg agt act      862
Gly Val Val Lys Leu Ala Asp Phe Gly Met Ala Lys His Leu Ser Thr
    195                 200                 205

gca gct cct aat ctt tca ttg aag gga aca cca tac tgg atg gct cct      910
Ala Ala Pro Asn Leu Ser Leu Lys Gly Thr Pro Tyr Trp Met Ala Pro
210                 215                 220                 225

gag gtt gtt cgg gct aca ctc gat aaa agt gca ggt tat gat ctt gct      958
Glu Val Val Arg Ala Thr Leu Asp Lys Ser Ala Gly Tyr Asp Leu Ala
                230                 235                 240

gtt gat att tgg agc ctt ggc tgc aca atc att gag atg ttt aca gga     1006
Val Asp Ile Trp Ser Leu Gly Cys Thr Ile Ile Glu Met Phe Thr Gly
            245                 250                 255

aag cca cct tgg agt ggt ctt gaa ggg ccg gct gca atg ttt aag gtg     1054
Lys Pro Pro Trp Ser Gly Leu Glu Gly Pro Ala Ala Met Phe Lys Val
        260                 265                 270

ttg cgc acg gac cca cca att cca gat aac ctg tcc cct gaa ggg aag     1102
Leu Arg Thr Asp Pro Pro Ile Pro Asp Asn Leu Ser Pro Glu Gly Lys
    275                 280                 285

gat ttt ctg aga tgc tgc ttc aag aga aac cct act gag cga cca aca     1150
Asp Phe Leu Arg Cys Cys Phe Lys Arg Asn Pro Thr Glu Arg Pro Thr
290                 295                 300                 305

gca aac aag ttg ctc gaa cac cca ttt atc caa act ttg aac cac tac     1198
Ala Asn Lys Leu Leu Glu His Pro Phe Ile Gln Thr Leu Asn His Tyr
                310                 315                 320

agc cca cat agc gtt gta cat tca ttt tcg gga att aaa tct cca gat     1246
Ser Pro His Ser Val Val His Ser Phe Ser Gly Ile Lys Ser Pro Asp
            325                 330                 335

act gtg cac agc agt gca aga gat aaa gta ccc tgg aaa agt gat tca     1294
Thr Val His Ser Ser Ala Arg Asp Lys Val Pro Trp Lys Ser Asp Ser
        340                 345                 350
```

```
tat atg aga ggg aaa cat gca aat ggt gaa aca agc agt gct cgg tcg      1342
Tyr Met Arg Gly Lys His Ala Asn Gly Glu Thr Ser Ser Ala Arg Ser
    355                 360                 365

tca gag tca ttg gct tat cgg ttg atg aca ccc cta ccc aac ttg gga      1390
Ser Glu Ser Leu Ala Tyr Arg Leu Met Thr Pro Leu Pro Asn Leu Gly
370                 375                 380                 385

acc aat agt ttg tca cct cca cct ttg agt tcg gcc tcc agt tct ggc      1438
Thr Asn Ser Leu Ser Pro Pro Pro Leu Ser Ser Ala Ser Ser Ser Gly
                390                 395                 400

tct gca gca cat aca ccg aac agt atg cat ttt tca gtc cca tac cat      1486
Ser Ala Ala His Thr Pro Asn Ser Met His Phe Ser Val Pro Tyr His
            405                 410                 415

cag cct agt cct ttg cca aag cct aac ggg aag gaa gca ata aat ttg      1534
Gln Pro Ser Pro Leu Pro Lys Pro Asn Gly Lys Glu Ala Ile Asn Leu
        420                 425                 430

gtt tcg cat tgaataacag catgccacct atcacaacat ctgggcaaag              1583
Val Ser His
    435

tccacataaa cagcgcaagt gtatccgtcc ctagcgtctg acccttgtac tgccagtaat    1643
agctaacaat tttctgggtc gtgtaaaaat aatttccatg taaagagatc accgcgatat    1703
gatcacaggg aaatgctgct aaatgaggga gctttgtgaa atctgttgta ccgtttgatt    1763
tctgggaaac tgattggatt gcccatgtaa catattttta ttagggtgta gctgtaggtg    1823
aaactaatgg tcaagattgt gccgacaaat aggacattag ggttgtagct gtagcgtgtt    1883
gttgatgctg tgtaatgatc atacttgatg taatcattat tgatttttt tggggaaggt    1943
aataatgcac gtcaattggc aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa    2003
aaaaaaaaaa aaaaaagggc ggccgc                                        2029

      <210> 2
      <211> 436
      <212> PRT
      <213> Zea mays

      <400> 2
Met His Pro Lys Leu Phe Pro Glu Asn His Val Ser Arg Pro Asp Gly
 1               5                  10                  15
Asn Gly Ser Ala Ser Phe His Pro Leu Pro Leu Pro Pro Ala Ser Val
            20                  25                  30
Ser Pro Lys Gln Thr Asn Ala Ser His Gln Leu Val Pro Lys Ala Glu
        35                  40                  45
Met Pro Ser Val Ala Gly Gln Trp Gln Lys Gly Lys Leu Leu Gly Ser
    50                  55                  60
Gly Thr Phe Gly Cys Val Tyr Glu Ala Thr Asn Arg His Thr Gly Ala
65                  70                  75                  80
Leu Cys Ala Met Lys Glu Val Asn Ile Ile Pro Asp Asp Ala Lys Ser
                85                  90                  95
Val Glu Ser Leu Lys Gln Leu Glu Gln Glu Ile Lys Phe Leu Ser Gln
            100                 105                 110
Phe Lys His Glu Asn Ile Val Gln Tyr Tyr Gly Ser Glu Thr Ile Glu
        115                 120                 125
Asp Arg Phe Tyr Ile Tyr Leu Glu Tyr Val His Pro Gly Ser Ile His
    130                 135                 140
Lys Tyr Val His Gln His Cys Gly Ser Leu Thr Glu Ala Val Ile Arg
145                 150                 155                 160
Asn Phe Thr Arg His Ile Leu Lys Gly Leu Ala Phe Leu His Ser Gln
                165                 170                 175
```

```
Lys Ile Met His Arg Asp Ile Lys Gly Ala Asn Leu Leu Val Asp Ile
            180                 185                 190
Asn Gly Val Val Lys Leu Ala Asp Phe Gly Met Ala Lys His Leu Ser
        195                 200                 205
Thr Ala Ala Pro Asn Leu Ser Leu Lys Gly Thr Pro Tyr Trp Met Ala
    210                 215                 220
Pro Glu Val Val Arg Ala Thr Leu Asp Lys Ser Ala Gly Tyr Asp Leu
225                 230                 235                 240
Ala Val Asp Ile Trp Ser Leu Gly Cys Thr Ile Ile Glu Met Phe Thr
                245                 250                 255
Gly Lys Pro Pro Trp Ser Gly Leu Glu Gly Pro Ala Ala Met Phe Lys
            260                 265                 270
Val Leu Arg Thr Asp Pro Pro Ile Pro Asp Asn Leu Ser Pro Glu Gly
        275                 280                 285
Lys Asp Phe Leu Arg Cys Cys Phe Lys Arg Asn Pro Thr Glu Arg Pro
    290                 295                 300
Thr Ala Asn Lys Leu Leu Glu His Pro Phe Ile Gln Thr Leu Asn His
305                 310                 315                 320
Tyr Ser Pro His Ser Val Val His Ser Phe Ser Gly Ile Lys Ser Pro
                325                 330                 335
Asp Thr Val His Ser Ser Ala Arg Asp Lys Val Pro Trp Lys Ser Asp
            340                 345                 350
Ser Tyr Met Arg Gly Lys His Ala Asn Gly Glu Thr Ser Ser Ala Arg
        355                 360                 365
Ser Ser Glu Ser Leu Ala Tyr Arg Leu Met Thr Pro Leu Pro Asn Leu
    370                 375                 380
Gly Thr Asn Ser Leu Ser Pro Pro Pro Leu Ser Ser Ala Ser Ser Ser
385                 390                 395                 400
Gly Ser Ala Ala His Thr Pro Asn Ser Met His Phe Ser Val Pro Tyr
                405                 410                 415
His Gln Pro Ser Pro Leu Pro Lys Pro Asn Gly Lys Glu Ala Ile Asn
            420                 425                 430
Leu Val Ser His
        435
```

<210> 3
<211> 20
<212> DNA
<213> Zea mays

<400> 3
atgcatccaa agttgttccc 20

<210> 4
<211> 20
<212> DNA
<213> Zea mays

<400> 4
tcaatgcgaa accaaattta 20

<210> 5
<211> 2063
<212> DNA
<213> Zea mays

<220>
<221> CDS
<222> (241)...(1713)

<400> 5
gtcgacccac gcgtccgggc ttctggtgtg ggtgttctga aacaatacga cttctgccca 60

```
cgcgccgccc ccccacactc cgcccgtgct gtccatttgc ttgaatatat tctttttcgt    120
tgtcatgtgt gccgtgggaa ttgctccatt cgattggtaa gtgctcgtct ctgaccatgc    180
agcctctcat ccggattgga ttggttccgg ctgttgcttg ctccaggaag aacactggac    240
atg gac ttc ttc acg gaa tat ggg gag gcg agt cag tat caa atc caa       288
Met Asp Phe Phe Thr Glu Tyr Gly Glu Ala Ser Gln Tyr Gln Ile Gln
 1               5                  10                  15

gag gtc att ggt aag gga agt tat gga gta gtt gct gct gca att gat      336
Glu Val Ile Gly Lys Gly Ser Tyr Gly Val Val Ala Ala Ala Ile Asp
            20                  25                  30

agt cac act ggg gag cgg gtt gcg atc aag aag ata aat gat att ttt      384
Ser His Thr Gly Glu Arg Val Ala Ile Lys Lys Ile Asn Asp Ile Phe
        35                  40                  45

ggt aat gtc tcg gat gct gct cgc att ttg cgg gaa atc aag ctt ctt      432
Gly Asn Val Ser Asp Ala Ala Arg Ile Leu Arg Glu Ile Lys Leu Leu
    50                  55                  60

cgt ctg ctt cgt cac cca aac ata gtc cag att aag cac att atg ttg      480
Arg Leu Leu Arg His Pro Asn Ile Val Gln Ile Lys His Ile Met Leu
65                  70                  75                  80

ccc cct acc cga agg gag ttc aaa gat atc tat gtt gtt ttt gag ctc      528
Pro Pro Thr Arg Arg Glu Phe Lys Asp Ile Tyr Val Val Phe Glu Leu
                85                  90                  95

atg gag tct gac ctg cat caa gtg atc aaa gca aat gat aac ctg act      576
Met Glu Ser Asp Leu His Gln Val Ile Lys Ala Asn Asp Asn Leu Thr
            100                 105                 110

cca gag cat cac cgc ttt ttc ctg tat caa ctt att cgt gct ctc aag      624
Pro Glu His His Arg Phe Phe Leu Tyr Gln Leu Ile Arg Ala Leu Lys
        115                 120                 125

tac atg cat tca gcc cat gta ttt cat cgt gat tta aag ccc agg aac      672
Tyr Met His Ser Ala His Val Phe His Arg Asp Leu Lys Pro Arg Asn
    130                 135                 140

ata ctg gcc aat tca gac agc aaa cta aag atc tgc gac ttt gga ctt      720
Ile Leu Ala Asn Ser Asp Ser Lys Leu Lys Ile Cys Asp Phe Gly Leu
145                 150                 155                 160

gca cgt gca tca ttt aac gac tct ctt tca gct atc tat tgg aca gat      768
Ala Arg Ala Ser Phe Asn Asp Ser Leu Ser Ala Ile Tyr Trp Thr Asp
                165                 170                 175

tat gtc gca aca agg tgg tac cgg gct cct gaa tta tgt ggc tcc ttt      816
Tyr Val Ala Thr Arg Trp Tyr Arg Ala Pro Glu Leu Cys Gly Ser Phe
            180                 185                 190

ttc tcc agt tac acc cct gca att gat att tgg agt ata ggg tgc att      864
Phe Ser Ser Tyr Thr Pro Ala Ile Asp Ile Trp Ser Ile Gly Cys Ile
        195                 200                 205

ttt gct gaa gtt ctc acc gga acg cca ttg ttt cct ggg agg aat gtt      912
Phe Ala Glu Val Leu Thr Gly Thr Pro Leu Phe Pro Gly Arg Asn Val
    210                 215                 220

gtg cac caa tta gac tta ata aca gat ctc ctt gga aca cca tcg ttt      960
Val His Gln Leu Asp Leu Ile Thr Asp Leu Leu Gly Thr Pro Ser Phe
225                 230                 235                 240
```

```
aga tcc tta tcc cag att cac agt gat aag gcc agg gaa tat ttg ctt     1008
Arg Ser Leu Ser Gln Ile His Ser Asp Lys Ala Arg Glu Tyr Leu Leu
                245                 250                 255

ggc atg ccg agg aag cgc cct att cca ttt tct cat aag ttc cgt aat     1056
Gly Met Pro Arg Lys Arg Pro Ile Pro Phe Ser His Lys Phe Arg Asn
            260                 265                 270

gcc gat cct ctg gct ctc cgg ttg cta gag cgt tta ctt gca ttt gac     1104
Ala Asp Pro Leu Ala Leu Arg Leu Leu Glu Arg Leu Leu Ala Phe Asp
        275                 280                 285

ccg aaa gat cga cct act gct gaa gag gct tta gct gat cca tac ttc     1152
Pro Lys Asp Arg Pro Thr Ala Glu Glu Ala Leu Ala Asp Pro Tyr Phe
    290                 295                 300

agt ggc ctt tct aaa ttg gaa ctt gag cct tca gca caa cca att ttg     1200
Ser Gly Leu Ser Lys Leu Glu Leu Glu Pro Ser Ala Gln Pro Ile Leu
305                 310                 315                 320

aag gtt gac ttt gaa ttc gag gga agg aaa cta aca aag gct ggt gtg     1248
Lys Val Asp Phe Glu Phe Glu Gly Arg Lys Leu Thr Lys Ala Gly Val
                325                 330                 335

aga gaa atg ata tat cgg gag att ttg gag tac cat ccc cag atg ctt     1296
Arg Glu Met Ile Tyr Arg Glu Ile Leu Glu Tyr His Pro Gln Met Leu
            340                 345                 350

cag gag tac atc gaa ggt gga gat cag att cac ttc cta tac cca agt     1344
Gln Glu Tyr Ile Glu Gly Gly Asp Gln Ile His Phe Leu Tyr Pro Ser
        355                 360                 365

ggg gtt gat cgc ttc cag cga cag ttt gct cat ctc gag gag aac tac     1392
Gly Val Asp Arg Phe Gln Arg Gln Phe Ala His Leu Glu Glu Asn Tyr
    370                 375                 380

agg aga gga gta gct agc acc ccc ccg cga agg cag ccc aca tct ctg     1440
Arg Arg Gly Val Ala Ser Thr Pro Pro Arg Arg Gln Pro Thr Ser Leu
385                 390                 395                 400

ccg agg gag cga gtc tgc tca tca gaa gat ggc cac agt cag gac tct     1488
Pro Arg Glu Arg Val Cys Ser Ser Glu Asp Gly His Ser Gln Asp Ser
                405                 410                 415

gac agc gaa gag cag aga gca gca tcc tac gtt gcc cga aca acc ata     1536
Asp Ser Glu Glu Gln Arg Ala Ala Ser Tyr Val Ala Arg Thr Thr Ile
            420                 425                 430

agc cct cca aga tca caa gag gag ggc ggc aag ctc cag cct gcc tac     1584
Ser Pro Pro Arg Ser Gln Glu Glu Gly Gly Lys Leu Gln Pro Ala Tyr
        435                 440                 445

cag agc gcc gcc gcc gcc gca gac tcc tgc gcc aag agc tat ctg aag     1632
Gln Ser Ala Ala Ala Ala Ala Asp Ser Cys Ala Lys Ser Tyr Leu Lys
    450                 455                 460

ggt gcc gct agt atc agt agt gct tcc agg cgt acc atc aag ggg gat     1680
Gly Ala Ala Ser Ile Ser Ser Ala Ser Arg Arg Thr Ile Lys Gly Asp
465                 470                 475                 480

aat ggc cga aag gag aag gga tct ccg agg ctg tgaaagaggt ggcgcaggct   1733
```

```
Asn Gly Arg Lys Glu Lys Gly Ser Pro Arg Leu
                485                 490

cgtgacctgt taggcaaaag aggtagcaac ggcgccaatg cgactgcaac gtgagatgtc   1793
ttagtcatca cgcatccaac aatgcaggag cacggagcac cacctcaccc ccccccccc    1853
ccccccagta caagtatgca ggtcatccta gctaggtcgc ccaaattaaa gcagcctgct   1913
tggaggaatt agatttgtgt actgttaagt ttttttttt ggaattcaac tgctagattc   1973
tggcagatct gtcaaccgca ctcgccttgg tttcccaaaa tttcctggtt tgaacaagtt   2033
taaaaaaaaa aaaaaaaaaa gggcggccgc                                    2063

      <210> 6
      <211> 491
      <212> PRT
      <213> Zea mays

      <400> 6
Met Asp Phe Phe Thr Glu Tyr Gly Glu Ala Ser Gln Tyr Gln Ile Gln
 1               5                  10                  15
Glu Val Ile Gly Lys Gly Ser Tyr Gly Val Val Ala Ala Ala Ile Asp
            20                  25                  30
Ser His Thr Gly Glu Arg Val Ala Ile Lys Lys Ile Asn Asp Ile Phe
        35                  40                  45
Gly Asn Val Ser Asp Ala Ala Arg Ile Leu Arg Glu Ile Lys Leu Leu
    50                  55                  60
Arg Leu Leu Arg His Pro Asn Ile Val Gln Ile Lys His Ile Met Leu
65                  70                  75                  80
Pro Pro Thr Arg Arg Glu Phe Lys Asp Ile Tyr Val Val Phe Glu Leu
                85                  90                  95
Met Glu Ser Asp Leu His Gln Val Ile Lys Ala Asn Asp Asn Leu Thr
            100                 105                 110
Pro Glu His His Arg Phe Phe Leu Tyr Gln Leu Ile Arg Ala Leu Lys
        115                 120                 125
Tyr Met His Ser Ala His Val Phe His Arg Asp Leu Lys Pro Arg Asn
    130                 135                 140
Ile Leu Ala Asn Ser Asp Ser Lys Leu Lys Ile Cys Asp Phe Gly Leu
145                 150                 155                 160
Ala Arg Ala Ser Phe Asn Asp Ser Leu Ser Ala Ile Tyr Trp Thr Asp
                165                 170                 175
Tyr Val Ala Thr Arg Trp Tyr Arg Ala Pro Glu Leu Cys Gly Ser Phe
            180                 185                 190
Phe Ser Ser Tyr Thr Pro Ala Ile Asp Ile Trp Ser Ile Gly Cys Ile
        195                 200                 205
Phe Ala Glu Val Leu Thr Gly Thr Pro Leu Phe Pro Gly Arg Asn Val
    210                 215                 220
Val His Gln Leu Asp Leu Ile Thr Asp Leu Leu Gly Thr Pro Ser Phe
225                 230                 235                 240
Arg Ser Leu Ser Gln Ile His Ser Asp Lys Ala Arg Glu Tyr Leu Leu
                245                 250                 255
Gly Met Pro Arg Lys Arg Pro Ile Pro Phe Ser His Lys Phe Arg Asn
            260                 265                 270
Ala Asp Pro Leu Ala Leu Arg Leu Leu Glu Arg Leu Leu Ala Phe Asp
        275                 280                 285
Pro Lys Asp Arg Pro Thr Ala Glu Glu Ala Leu Ala Asp Pro Tyr Phe
    290                 295                 300
Ser Gly Leu Ser Lys Leu Glu Leu Glu Pro Ser Ala Gln Pro Ile Leu
305                 310                 315                 320
Lys Val Asp Phe Glu Phe Glu Gly Arg Lys Leu Thr Lys Ala Gly Val
                325                 330                 335
Arg Glu Met Ile Tyr Arg Glu Ile Leu Glu Tyr His Pro Gln Met Leu
            340                 345                 350
Gln Glu Tyr Ile Glu Gly Gly Asp Gln Ile His Phe Leu Tyr Pro Ser
        355                 360                 365
```

```
Gly Val Asp Arg Phe Gln Arg Gln Phe Ala His Leu Glu Glu Asn Tyr
    370                 375                 380
Arg Arg Gly Val Ala Ser Thr Pro Pro Arg Arg Gln Pro Thr Ser Leu
385                 390                 395                 400
Pro Arg Glu Arg Val Cys Ser Ser Glu Asp Gly His Ser Gln Asp Ser
                405                 410                 415
Asp Ser Glu Glu Gln Arg Ala Ala Ser Tyr Val Ala Arg Thr Thr Ile
            420                 425                 430
Ser Pro Pro Arg Ser Gln Glu Glu Gly Gly Lys Leu Gln Pro Ala Tyr
        435                 440                 445
Gln Ser Ala Ala Ala Ala Ala Asp Ser Cys Ala Lys Ser Tyr Leu Lys
    450                 455                 460
Gly Ala Ala Ser Ile Ser Ser Ala Ser Arg Arg Thr Ile Lys Gly Asp
465                 470                 475                 480
Asn Gly Arg Lys Glu Lys Gly Ser Pro Arg Leu
                485                 490
```

<210> 7
<211> 20
<212> DNA
<213> Zea mays

<400> 7
atggacttct tcacggaata 20

<210> 8
<211> 20
<212> DNA
<213> Zea mays

<400> 8
tcacagcctc ggagatccct 20

<210> 9
<211> 753
<212> DNA
<213> Zea mays

<220>
<221> CDS
<222> (1)...(750)

<400> 9

```
atg aag gga aaa ggg ggg cct gac aat act caa tgt gga tac cgt gga      48
Met Lys Gly Lys Gly Gly Pro Asp Asn Thr Gln Cys Gly Tyr Arg Gly
 1               5                  10                  15

gtg agg cag cgt act tgg ggg aag tgg gtt gct gaa ata aga gag cca      96
Val Arg Gln Arg Thr Trp Gly Lys Trp Val Ala Glu Ile Arg Glu Pro
            20                  25                  30

aat cgt gtc gac aga ctc tgg ctg ggt acc ttc cca acc gcg gag gat     144
Asn Arg Val Asp Arg Leu Trp Leu Gly Thr Phe Pro Thr Ala Glu Asp
        35                  40                  45

gca gct agg gcc tat gat gag gca gcc aga gcg atg tat gga gac ttg     192
Ala Ala Arg Ala Tyr Asp Glu Ala Ala Arg Ala Met Tyr Gly Asp Leu
    50                  55                  60

gca cgg act aac ttc ccc gga cag gat gca aca acc tct gcc caa gct     240
Ala Arg Thr Asn Phe Pro Gly Gln Asp Ala Thr Thr Ser Ala Gln Ala
65                  70                  75                  80
```

```
gct cta gca tcg acc tct gcc cag gct gat cca aca gct gtc gaa gct      288
Ala Leu Ala Ser Thr Ser Ala Gln Ala Asp Pro Thr Ala Val Glu Ala
                 85                  90                  95

ctt cag act ggc acg tca tgc gag tcg aca acg aca tca aat tac tcg      336
Leu Gln Thr Gly Thr Ser Cys Glu Ser Thr Thr Thr Ser Asn Tyr Ser
            100                 105                 110

gac atc gca tcc acc tca cac aag cct gag cct gaa gcc tct gac atc      384
Asp Ile Ala Ser Thr Ser His Lys Pro Glu Pro Glu Ala Ser Asp Ile
        115                 120                 125

tcg agc tcc cta aag gca aaa tgt cca gct gga tca tgt ggt atc caa      432
Ser Ser Ser Leu Lys Ala Lys Cys Pro Ala Gly Ser Cys Gly Ile Gln
    130                 135                 140

gac ggt aca ccc agt gta gct gac aag gag gtc ttt ggg ccg ttg gag      480
Asp Gly Thr Pro Ser Val Ala Asp Lys Glu Val Phe Gly Pro Leu Glu
145                 150                 155                 160

cct atc aca aat ctt cca gat ggt ggt gat ggt ttt gat atc ggt gag      528
Pro Ile Thr Asn Leu Pro Asp Gly Gly Asp Gly Phe Asp Ile Gly Glu
                165                 170                 175

atg ctg agg atg atg gaa agc gat cca cat aat gca ggt gga gct gac      576
Met Leu Arg Met Met Glu Ser Asp Pro His Asn Ala Gly Gly Ala Asp
            180                 185                 190

gct ggc atg ggg cag ccc tgg tat ctt gat gag ctg gat tcg agt gtc      624
Ala Gly Met Gly Gln Pro Trp Tyr Leu Asp Glu Leu Asp Ser Ser Val
        195                 200                 205

ttg gag agc atg ctc cag cca gag cca gag cca gag cca gag cca ttc      672
Leu Glu Ser Met Leu Gln Pro Glu Pro Glu Pro Glu Pro Glu Pro Phe
    210                 215                 220

ctg atg tct gaa gaa ccg gac atg ttt ctt gct ggc ttc gaa agc gct      720
Leu Met Ser Glu Glu Pro Asp Met Phe Leu Ala Gly Phe Glu Ser Ala
225                 230                 235                 240

ggt ttc gtc gag ggt ctg gag cgg cta aac tga                          753
Gly Phe Val Glu Gly Leu Glu Arg Leu Asn
                245                 250
```

<210> 10
<211> 250
<212> PRT
<213> Zea mays

<400> 10

```
Met Lys Gly Lys Gly Gly Pro Asp Asn Thr Gln Cys Gly Tyr Arg Gly
 1               5                  10                  15
Val Arg Gln Arg Thr Trp Gly Lys Trp Val Ala Glu Ile Arg Glu Pro
            20                  25                  30
Asn Arg Val Asp Arg Leu Trp Leu Gly Thr Phe Pro Thr Ala Glu Asp
        35                  40                  45
Ala Ala Arg Ala Tyr Asp Glu Ala Ala Arg Ala Met Tyr Gly Asp Leu
    50                  55                  60
Ala Arg Thr Asn Phe Pro Gly Gln Asp Ala Thr Thr Ser Ala Gln Ala
65                  70                  75                  80
```

```
Ala Leu Ala Ser Thr Ser Ala Gln Ala Asp Pro Thr Ala Val Glu Ala
                85                  90                  95
Leu Gln Thr Gly Thr Ser Cys Glu Ser Thr Thr Thr Ser Asn Tyr Ser
            100                 105                 110
Asp Ile Ala Ser Thr Ser His Lys Pro Glu Pro Glu Ala Ser Asp Ile
        115                 120                 125
Ser Ser Ser Leu Lys Ala Lys Cys Pro Ala Gly Ser Cys Gly Ile Gln
    130                 135                 140
Asp Gly Thr Pro Ser Val Ala Asp Lys Glu Val Phe Gly Pro Leu Glu
145                 150                 155                 160
Pro Ile Thr Asn Leu Pro Asp Gly Gly Asp Gly Phe Asp Ile Gly Glu
                165                 170                 175
Met Leu Arg Met Met Glu Ser Asp Pro His Asn Ala Gly Gly Ala Asp
            180                 185                 190
Ala Gly Met Gly Gln Pro Trp Tyr Leu Asp Glu Leu Asp Ser Ser Val
        195                 200                 205
Leu Glu Ser Met Leu Gln Pro Glu Pro Glu Pro Glu Pro Glu Pro Phe
    210                 215                 220
Leu Met Ser Glu Glu Pro Asp Met Phe Leu Ala Gly Phe Glu Ser Ala
225                 230                 235                 240
Gly Phe Val Glu Gly Leu Glu Arg Leu Asn
                245                 250
```

<210> 11
<211> 20
<212> DNA
<213> Zea mays

<400> 11
atgaagggaa aaggggggcc 20

<210> 12
<211> 20
<212> DNA
<213> Zea mays

<400> 12
tcagtttagc cgctccagac 20

<210> 13
<211> 1110
<212> DNA
<213> Zea mays

<220>
<221> CDS
<222> (1)...(1107)

<400> 13

```
atg gcg atg atg gtg gat ccc ccc aat ggc atg gca agt caa gga aag      48
Met Ala Met Met Val Asp Pro Pro Asn Gly Met Ala Ser Gln Gly Lys
 1               5                  10                  15

cat tac tat act atg tgg cag aca ctg ttt gag atc gac acc aag tac      96
His Tyr Tyr Thr Met Trp Gln Thr Leu Phe Glu Ile Asp Thr Lys Tyr
            20                  25                  30

gtg ccg atc aag ccc atc gga aga gga gct tac ggt gtc gtt tgc tca     144
Val Pro Ile Lys Pro Ile Gly Arg Gly Ala Tyr Gly Val Val Cys Ser
        35                  40                  45

tct gtt aac cgc gag acc aat gaa aaa gtc gcg ata aaa aag ata aac     192
```

```
Ser Val Asn Arg Glu Thr Asn Glu Lys Val Ala Ile Lys Lys Ile Asn
     50                  55                  60

aac gtc ttt gac aac cgt gtg gat gcg cta agg acg ctg aga gag ctg      240
Asn Val Phe Asp Asn Arg Val Asp Ala Leu Arg Thr Leu Arg Glu Leu
 65                  70                  75                  80

aaa ctc ctt cgg cac ctg agg cac gag aat gtt att gct ctc aag gac      288
Lys Leu Leu Arg His Leu Arg His Glu Asn Val Ile Ala Leu Lys Asp
                 85                  90                  95

ata atg atg cct gcg cat aga agg agc ttc aag gac gtt tac ttg gtt      336
Ile Met Met Pro Ala His Arg Arg Ser Phe Lys Asp Val Tyr Leu Val
            100                 105                 110

tat gag ctc atg gac act gat ctg cat caa ata att aag tca tct caa      384
Tyr Glu Leu Met Asp Thr Asp Leu His Gln Ile Ile Lys Ser Ser Gln
        115                 120                 125

cca cta tcc aat gac cac tgc cag tat ttc ctt ttt cag ctg ctc cga      432
Pro Leu Ser Asn Asp His Cys Gln Tyr Phe Leu Phe Gln Leu Leu Arg
    130                 135                 140

ggc ctg aag tac ctc cat tca gcc ggg ata ctc cac aga gac cta aag      480
Gly Leu Lys Tyr Leu His Ser Ala Gly Ile Leu His Arg Asp Leu Lys
145                 150                 155                 160

cca ggg aac ctc ctg gtc aac gca aac tgt gac ctg aag ata tgc gac      528
Pro Gly Asn Leu Leu Val Asn Ala Asn Cys Asp Leu Lys Ile Cys Asp
                165                 170                 175

ttc ggg ctc gcc cgc acg aac aac acc aag ggc cag ttc atg acg gag      576
Phe Gly Leu Ala Arg Thr Asn Asn Thr Lys Gly Gln Phe Met Thr Glu
            180                 185                 190

tac gtg gtg acc cgc tgg tac agg gca ccc gag ctg ctg ctc tgc tgc      624
Tyr Val Val Thr Arg Trp Tyr Arg Ala Pro Glu Leu Leu Leu Cys Cys
        195                 200                 205

gac aac tac ggc acg tcc att gac gtc tgg tct gtg ggg tgc ata ttc      672
Asp Asn Tyr Gly Thr Ser Ile Asp Val Trp Ser Val Gly Cys Ile Phe
    210                 215                 220

gcg gag ctg ctt ggc cgc aag ccg atc ttc cca gga acc gag tgc ctg      720
Ala Glu Leu Leu Gly Arg Lys Pro Ile Phe Pro Gly Thr Glu Cys Leu
225                 230                 235                 240

aac cag ctc aag ctc atc gtc aac gtc ctc ggc acc atg ggc gag gcc      768
Asn Gln Leu Lys Leu Ile Val Asn Val Leu Gly Thr Met Gly Glu Ala
                245                 250                 255

gac ctc gcg ttc atc gac aac ccg aag gcc cgc aag tac atc aag tcc      816
Asp Leu Ala Phe Ile Asp Asn Pro Lys Ala Arg Lys Tyr Ile Lys Ser
            260                 265                 270

ctt ccg tac gcc ccg ggc gcc ccc ttc acc ggc atg tac cct cag gcg      864
Leu Pro Tyr Ala Pro Gly Ala Pro Phe Thr Gly Met Tyr Pro Gln Ala
        275                 280                 285

cac cct ctc gcc atc gac ctg ctg cag aag atg ctc gtg ttc gac ccg      912
His Pro Leu Ala Ile Asp Leu Leu Gln Lys Met Leu Val Phe Asp Pro
    290                 295                 300
```

```
tcc aag agg atc agc gtc acc gag gcg ctg gag cac ccg tac atg tct      960
Ser Lys Arg Ile Ser Val Thr Glu Ala Leu Glu His Pro Tyr Met Ser
305                 310                 315                 320

ccg ctc tat gac ccg agc gcg aac cct ccc gcg cag gtg ccg atc gac     1008
Pro Leu Tyr Asp Pro Ser Ala Asn Pro Pro Ala Gln Val Pro Ile Asp
                325                 330                 335

ctc gac atc gac gag aac ctc ggc gtc gac atg atc agg gag atg atg     1056
Leu Asp Ile Asp Glu Asn Leu Gly Val Asp Met Ile Arg Glu Met Met
            340                 345                 350

tgg cag gag atg atc cac tac cac ccc gag gtc ctc acg aga atc agc     1104
Trp Gln Glu Met Ile His Tyr His Pro Glu Val Leu Thr Arg Ile Ser
        355                 360                 365

atg tga                                                            1110
Met
```

<210> 14
<211> 369
<212> PRT
<213> Zea mays

<400> 14

```
Met Ala Met Met Val Asp Pro Pro Asn Gly Met Ala Ser Gln Gly Lys
 1               5                  10                  15
His Tyr Tyr Thr Met Trp Gln Thr Leu Phe Glu Ile Asp Thr Lys Tyr
            20                  25                  30
Val Pro Ile Lys Pro Ile Gly Arg Gly Ala Tyr Gly Val Val Cys Ser
        35                  40                  45
Ser Val Asn Arg Glu Thr Asn Glu Lys Val Ala Ile Lys Lys Ile Asn
    50                  55                  60
Asn Val Phe Asp Asn Arg Val Asp Ala Leu Arg Thr Leu Arg Glu Leu
65                  70                  75                  80
Lys Leu Leu Arg His Leu Arg His Glu Asn Val Ile Ala Leu Lys Asp
                85                  90                  95
Ile Met Met Pro Ala His Arg Arg Ser Phe Lys Asp Val Tyr Leu Val
            100                 105                 110
Tyr Glu Leu Met Asp Thr Asp Leu His Gln Ile Ile Lys Ser Ser Gln
        115                 120                 125
Pro Leu Ser Asn Asp His Cys Gln Tyr Phe Leu Phe Gln Leu Leu Arg
    130                 135                 140
Gly Leu Lys Tyr Leu His Ser Ala Gly Ile Leu His Arg Asp Leu Lys
145                 150                 155                 160
Pro Gly Asn Leu Leu Val Asn Ala Asn Cys Asp Leu Lys Ile Cys Asp
                165                 170                 175
Phe Gly Leu Ala Arg Thr Asn Asn Thr Lys Gly Gln Phe Met Thr Glu
            180                 185                 190
Tyr Val Val Thr Arg Trp Tyr Arg Ala Pro Glu Leu Leu Leu Cys Cys
        195                 200                 205
Asp Asn Tyr Gly Thr Ser Ile Asp Val Trp Ser Val Gly Cys Ile Phe
    210                 215                 220
Ala Glu Leu Leu Gly Arg Lys Pro Ile Phe Pro Gly Thr Glu Cys Leu
225                 230                 235                 240
Asn Gln Leu Lys Leu Ile Val Asn Val Leu Gly Thr Met Gly Glu Ala
                245                 250                 255
Asp Leu Ala Phe Ile Asp Asn Pro Lys Ala Arg Lys Tyr Ile Lys Ser
            260                 265                 270
```

```
Leu Pro Tyr Ala Pro Gly Ala Pro Phe Thr Gly Met Tyr Pro Gln Ala
        275                 280                 285
His Pro Leu Ala Ile Asp Leu Leu Gln Lys Met Leu Val Phe Asp Pro
    290                 295                 300
Ser Lys Arg Ile Ser Val Thr Glu Ala Leu Glu His Pro Tyr Met Ser
305                 310                 315                 320
Pro Leu Tyr Asp Pro Ser Ala Asn Pro Pro Ala Gln Val Pro Ile Asp
                325                 330                 335
Leu Asp Ile Asp Glu Asn Leu Gly Val Asp Met Ile Arg Glu Met Met
            340                 345                 350
Trp Gln Glu Met Ile His Tyr His Pro Glu Val Leu Thr Arg Ile Ser
        355                 360                 365
Met
```

<210> 15
<211> 20
<212> DNA
<213> Zea mays

<400> 15
atggcgatga tggtggatcc 20

<210> 16
<211> 20
<212> DNA
<213> Zea mays

<400> 16
tcacatgctg attctcgtga 20

<210> 17
<211> 1119
<212> DNA
<213> Zea mays

<220>
<221> CDS
<222> (1)...(1116)

<400> 17

```
atg gat tcc tcc ggc ggc ggc ggt gga gcg cag atc aag ggt atg gcg       48
Met Asp Ser Ser Gly Gly Gly Gly Gly Ala Gln Ile Lys Gly Met Ala
 1               5                  10                  15

acg cac ggg ggc cgc tac gtg ctg tac aac gtg tac gga aac ctc ttc       96
Thr His Gly Gly Arg Tyr Val Leu Tyr Asn Val Tyr Gly Asn Leu Phe
            20                  25                  30

gag gtc tcc tcc aag tac gcc cca ccc atc cgc ccc atc ggt cgc ggc      144
Glu Val Ser Ser Lys Tyr Ala Pro Pro Ile Arg Pro Ile Gly Arg Gly
        35                  40                  45

gcc tac ggc att gtc tgc gcg gct gtc aac tcg cag tca ggg gag gag      192
Ala Tyr Gly Ile Val Cys Ala Ala Val Asn Ser Gln Ser Gly Glu Glu
    50                  55                  60

gtt gcg atc aag aag gtt ggc aat gcg ttc gac aac cac atc gac gcc      240
Val Ala Ile Lys Lys Val Gly Asn Ala Phe Asp Asn His Ile Asp Ala
65                  70                  75                  80

aag cgg acg ctc agg gaa atc aag ctg ctg cgc cac atg gac cat gag      288
```

```
Lys Arg Thr Leu Arg Glu Ile Lys Leu Leu Arg His Met Asp His Glu
                85                  90                  95

aac atc ctt gcc tta aag gat gta att cgg ccc cca act aga gag aac      336
Asn Ile Leu Ala Leu Lys Asp Val Ile Arg Pro Pro Thr Arg Glu Asn
            100                 105                 110

ttt aat gac gtg tac att gtt act gag tta atg gat aca gat ctc cat      384
Phe Asn Asp Val Tyr Ile Val Thr Glu Leu Met Asp Thr Asp Leu His
        115                 120                 125

cag atc gta cgc tca aat cag cca ttg act gat gat cat tgc cag tac      432
Gln Ile Val Arg Ser Asn Gln Pro Leu Thr Asp Asp His Cys Gln Tyr
    130                 135                 140

ttc ttg tat cag ttg tta cga ggg cta aaa tat gtg cac tca gca aat      480
Phe Leu Tyr Gln Leu Leu Arg Gly Leu Lys Tyr Val His Ser Ala Asn
145                 150                 155                 160

ata ttg cac cgc gat ctg aag ccg agc aat ttg ttc cta aat gca aat      528
Ile Leu His Arg Asp Leu Lys Pro Ser Asn Leu Phe Leu Asn Ala Asn
                165                 170                 175

tgt gac ctc aag att gca gac ttt ggg ctt gca agg acc act tca gag      576
Cys Asp Leu Lys Ile Ala Asp Phe Gly Leu Ala Arg Thr Thr Ser Glu
            180                 185                 190

aca gat ctc atg aca gag tat gtg gtc act cgt tgg tac cgg gca cca      624
Thr Asp Leu Met Thr Glu Tyr Val Val Thr Arg Trp Tyr Arg Ala Pro
        195                 200                 205

gag ctg ctg ttg aac tgt tca cag tat act gct gcc att gat gtc tgg      672
Glu Leu Leu Leu Asn Cys Ser Gln Tyr Thr Ala Ala Ile Asp Val Trp
    210                 215                 220

tca gtt gga tgc ata cta ggt gaa atc gtt act cgt caa ccc ctg ttt      720
Ser Val Gly Cys Ile Leu Gly Glu Ile Val Thr Arg Gln Pro Leu Phe
225                 230                 235                 240

cct gga cgg gat tac atc cag caa tta aaa ttg atc act gag ctc ata      768
Pro Gly Arg Asp Tyr Ile Gln Gln Leu Lys Leu Ile Thr Glu Leu Ile
                245                 250                 255

ggc tct cca gat gat gca agc ctg gga ttt ctt cga agt gat aat gca      816
Gly Ser Pro Asp Asp Ala Ser Leu Gly Phe Leu Arg Ser Asp Asn Ala
            260                 265                 270

aaa aga tac atg aaa caa cta cca cag ttt cca aga cag gac ttc cgc      864
Lys Arg Tyr Met Lys Gln Leu Pro Gln Phe Pro Arg Gln Asp Phe Arg
        275                 280                 285

ctg cgt ttc cgc aac atg tct cct ggc gca gtc gat ttg ttg gaa agg      912
Leu Arg Phe Arg Asn Met Ser Pro Gly Ala Val Asp Leu Leu Glu Arg
    290                 295                 300

atg ctt gtg ttt gat cca agc aga cgg att aca gtt gat gag gct ctg      960
Met Leu Val Phe Asp Pro Ser Arg Arg Ile Thr Val Asp Glu Ala Leu
305                 310                 315                 320

cat cat cca tac ttg gct tca ctt cat gag atc aat gat gaa cct acc     1008
His His Pro Tyr Leu Ala Ser Leu His Glu Ile Asn Asp Glu Pro Thr
                325                 330                 335
```

```
tgc cct gca cct ttc agc ttt gat ttt gag caa cca tcc ttt aca gaa      1056
Cys Pro Ala Pro Phe Ser Phe Asp Phe Glu Gln Pro Ser Phe Thr Glu
            340                 345                 350

gcg cat ata aaa gaa ctc atc tgg agg gaa tct tta gca ttt aac cca      1104
Ala His Ile Lys Glu Leu Ile Trp Arg Glu Ser Leu Ala Phe Asn Pro
        355                 360                 365

gag cct ccc tac taa                                                 1119
Glu Pro Pro Tyr
    370
```

<210> 18
<211> 372
<212> PRT
<213> Zea mays

<400> 18

```
Met Asp Ser Ser Gly Gly Gly Gly Gly Ala Gln Ile Lys Gly Met Ala
 1               5                  10                  15
Thr His Gly Gly Arg Tyr Val Leu Tyr Asn Val Tyr Gly Asn Leu Phe
            20                  25                  30
Glu Val Ser Ser Lys Tyr Ala Pro Pro Ile Arg Pro Ile Gly Arg Gly
        35                  40                  45
Ala Tyr Gly Ile Val Cys Ala Ala Val Asn Ser Gln Ser Gly Glu Glu
    50                  55                  60
Val Ala Ile Lys Lys Val Gly Asn Ala Phe Asp Asn His Ile Asp Ala
65                  70                  75                  80
Lys Arg Thr Leu Arg Glu Ile Lys Leu Leu Arg His Met Asp His Glu
                85                  90                  95
Asn Ile Leu Ala Leu Lys Asp Val Ile Arg Pro Pro Thr Arg Glu Asn
            100                 105                 110
Phe Asn Asp Val Tyr Ile Val Thr Glu Leu Met Asp Thr Asp Leu His
        115                 120                 125
Gln Ile Val Arg Ser Asn Gln Pro Leu Thr Asp Asp His Cys Gln Tyr
    130                 135                 140
Phe Leu Tyr Gln Leu Leu Arg Gly Leu Lys Tyr Val His Ser Ala Asn
145                 150                 155                 160
Ile Leu His Arg Asp Leu Lys Pro Ser Asn Leu Phe Leu Asn Ala Asn
                165                 170                 175
Cys Asp Leu Lys Ile Ala Asp Phe Gly Leu Ala Arg Thr Thr Ser Glu
            180                 185                 190
Thr Asp Leu Met Thr Glu Tyr Val Val Thr Arg Trp Tyr Arg Ala Pro
        195                 200                 205
Glu Leu Leu Leu Asn Cys Ser Gln Tyr Thr Ala Ala Ile Asp Val Trp
    210                 215                 220
Ser Val Gly Cys Ile Leu Gly Glu Ile Val Thr Arg Gln Pro Leu Phe
225                 230                 235                 240
Pro Gly Arg Asp Tyr Ile Gln Gln Leu Lys Leu Ile Thr Glu Leu Ile
                245                 250                 255
Gly Ser Pro Asp Asp Ala Ser Leu Gly Phe Leu Arg Ser Asp Asn Ala
            260                 265                 270
Lys Arg Tyr Met Lys Gln Leu Pro Gln Phe Pro Arg Gln Asp Phe Arg
        275                 280                 285
Leu Arg Phe Arg Asn Met Ser Pro Gly Ala Val Asp Leu Leu Glu Arg
    290                 295                 300
Met Leu Val Phe Asp Pro Ser Arg Arg Ile Thr Val Asp Glu Ala Leu
305                 310                 315                 320
His His Pro Tyr Leu Ala Ser Leu His Glu Ile Asn Asp Glu Pro Thr
                325                 330                 335
```

```
Cys Pro Ala Pro Phe Ser Phe Asp Phe Glu Gln Pro Ser Phe Thr Glu
            340                 345                 350
Ala His Ile Lys Glu Leu Ile Trp Arg Glu Ser Leu Ala Phe Asn Pro
        355                 360                 365
Glu Pro Pro Tyr
    370
```

<210> 19
<211> 20
<212> DNA
<213> Zea mays

<400> 19
atggattcct ccggcggcgg 20

<210> 20
<211> 20
<212> DNA
<213> Zea mays

<400> 20
ttagtaggga ggctctgggt 20

<210> 21
<211> 1113
<212> DNA
<213> Zea mays

<220>
<221> CDS
<222> (1)...(1110)

<400> 21

```
atg gcg atg atg gtg gat cct ccg aat gga atc ggg aac caa gga aag      48
Met Ala Met Met Val Asp Pro Pro Asn Gly Ile Gly Asn Gln Gly Lys
 1               5                  10                  15

cat tac tac tct atg tgg cag acc tta ttt gag ata gac acc aaa tat      96
His Tyr Tyr Ser Met Trp Gln Thr Leu Phe Glu Ile Asp Thr Lys Tyr
            20                  25                  30

gtg ccg atc aag ccc att ggt cga gga gct tat ggg ata gtt tgt tca     144
Val Pro Ile Lys Pro Ile Gly Arg Gly Ala Tyr Gly Ile Val Cys Ser
        35                  40                  45

tcc att aat cgt gaa aca aat gag aaa gta gca ata aag aag ata cac     192
Ser Ile Asn Arg Glu Thr Asn Glu Lys Val Ala Ile Lys Lys Ile His
    50                  55                  60

aac gtt ttc gac aac cgt gtg gat gca cta cgg acc ttg cgg gag ctg     240
Asn Val Phe Asp Asn Arg Val Asp Ala Leu Arg Thr Leu Arg Glu Leu
65                  70                  75                  80

aaa ctc ctt cgc cat ctc cgg cat gag aat gtc att gct ttg aag gat     288
Lys Leu Leu Arg His Leu Arg His Glu Asn Val Ile Ala Leu Lys Asp
                85                  90                  95

ata atg atg cca ata cac agg aga agc ttt aag gat gtg tac ttg gta     336
Ile Met Met Pro Ile His Arg Arg Ser Phe Lys Asp Val Tyr Leu Val
            100                 105                 110

tac gaa ctc atg gat act gat ttg cac cag ata atc aaa tca cct cag     384
```

```
Tyr Glu Leu Met Asp Thr Asp Leu His Gln Ile Ile Lys Ser Pro Gln
        115                 120                 125

ggc ctt tcc aat gac cac tgc cag tat ttt ctt ttt cag ttg ctc cga      432
Gly Leu Ser Asn Asp His Cys Gln Tyr Phe Leu Phe Gln Leu Leu Arg
    130                 135                 140

gga ctc aaa tat ctc cat tca gca gaa ata ctc cac aga gac cta aaa      480
Gly Leu Lys Tyr Leu His Ser Ala Glu Ile Leu His Arg Asp Leu Lys
145                 150                 155                 160

cct gga aac ctg ctg gtg aat gca aat tgt gat ctg aag ata tgt gat      528
Pro Gly Asn Leu Leu Val Asn Ala Asn Cys Asp Leu Lys Ile Cys Asp
                165                 170                 175

ttt ggt ctc gca cgt aca aac agt agc aaa ggc cag ttc atg act gaa      576
Phe Gly Leu Ala Arg Thr Asn Ser Ser Lys Gly Gln Phe Met Thr Glu
            180                 185                 190

tac gtc gtc acc cgc tgg tac aga gct cct gag ctg ctc ctc tgc tgc      624
Tyr Val Val Thr Arg Trp Tyr Arg Ala Pro Glu Leu Leu Leu Cys Cys
        195                 200                 205

gac aac tac ggc aca tcc ata gac gtc tgg tct gtt ggg tgc atc ttt      672
Asp Asn Tyr Gly Thr Ser Ile Asp Val Trp Ser Val Gly Cys Ile Phe
    210                 215                 220

gct gag ctc ctt ggc cgc aag cca ata ttt cca gga act gaa tgc ctg      720
Ala Glu Leu Leu Gly Arg Lys Pro Ile Phe Pro Gly Thr Glu Cys Leu
225                 230                 235                 240

aat caa ctc aag ctc ata gtg aac gtc ctc ggc aca atg agt gag gct      768
Asn Gln Leu Lys Leu Ile Val Asn Val Leu Gly Thr Met Ser Glu Ala
                245                 250                 255

gac cta gag ttc atc gac aac cca aag gct cgg aga tac atc aag tcc      816
Asp Leu Glu Phe Ile Asp Asn Pro Lys Ala Arg Arg Tyr Ile Lys Ser
            260                 265                 270

ctt ccc tat acc cct ggt gtt cct ctc gta agt atg tac cca cat gcg      864
Leu Pro Tyr Thr Pro Gly Val Pro Leu Val Ser Met Tyr Pro His Ala
        275                 280                 285

cac cct ctt gcc att gat ctg ttg cag aag atg ctc atc ttc gac ccc      912
His Pro Leu Ala Ile Asp Leu Leu Gln Lys Met Leu Ile Phe Asp Pro
    290                 295                 300

acc aaa agg atc agt gtc acc gag gct ctc gag cac cct tac atg tcc      960
Thr Lys Arg Ile Ser Val Thr Glu Ala Leu Glu His Pro Tyr Met Ser
305                 310                 315                 320

cct ctg tat gac cca agc gca aat ccc cca gcc caa gtg ccc atc gat     1008
Pro Leu Tyr Asp Pro Ser Ala Asn Pro Pro Ala Gln Val Pro Ile Asp
                325                 330                 335

ctg gac ata gac gaa aac atc agc tca gag atg atc cgg gaa atg atg     1056
Leu Asp Ile Asp Glu Asn Ile Ser Ser Glu Met Ile Arg Glu Met Met
            340                 345                 350

tgg cag gag atg ctt cac tac cac cct gaa gtt gcc aca gca ata agc     1104
Trp Gln Glu Met Leu His Tyr His Pro Glu Val Ala Thr Ala Ile Ser
        355                 360                 365
```

```
atg tca tga                                                          1113
Met Ser
    370
```

<210> 22
<211> 370
<212> PRT
<213> Zea mays

<400> 22

```
Met Ala Met Met Val Asp Pro Pro Asn Gly Ile Gly Asn Gln Gly Lys
 1               5                  10                  15
His Tyr Tyr Ser Met Trp Gln Thr Leu Phe Glu Ile Asp Thr Lys Tyr
            20                  25                  30
Val Pro Ile Lys Pro Ile Gly Arg Gly Ala Tyr Gly Ile Val Cys Ser
        35                  40                  45
Ser Ile Asn Arg Glu Thr Asn Glu Lys Val Ala Ile Lys Lys Ile His
    50                  55                  60
Asn Val Phe Asp Asn Arg Val Asp Ala Leu Arg Thr Leu Arg Glu Leu
65                  70                  75                  80
Lys Leu Leu Arg His Leu Arg His Glu Asn Val Ile Ala Leu Lys Asp
                85                  90                  95
Ile Met Met Pro Ile His Arg Arg Ser Phe Lys Asp Val Tyr Leu Val
            100                 105                 110
Tyr Glu Leu Met Asp Thr Asp Leu His Gln Ile Ile Lys Ser Pro Gln
        115                 120                 125
Gly Leu Ser Asn Asp His Cys Gln Tyr Phe Leu Phe Gln Leu Leu Arg
    130                 135                 140
Gly Leu Lys Tyr Leu His Ser Ala Glu Ile Leu His Arg Asp Leu Lys
145                 150                 155                 160
Pro Gly Asn Leu Leu Val Asn Ala Asn Cys Asp Leu Lys Ile Cys Asp
                165                 170                 175
Phe Gly Leu Ala Arg Thr Asn Ser Ser Lys Gly Gln Phe Met Thr Glu
            180                 185                 190
Tyr Val Val Thr Arg Trp Tyr Arg Ala Pro Glu Leu Leu Leu Cys Cys
        195                 200                 205
Asp Asn Tyr Gly Thr Ser Ile Asp Val Trp Ser Val Gly Cys Ile Phe
    210                 215                 220
Ala Glu Leu Leu Gly Arg Lys Pro Ile Phe Pro Gly Thr Glu Cys Leu
225                 230                 235                 240
Asn Gln Leu Lys Leu Ile Val Asn Val Leu Gly Thr Met Ser Glu Ala
                245                 250                 255
Asp Leu Glu Phe Ile Asp Asn Pro Lys Ala Arg Arg Tyr Ile Lys Ser
            260                 265                 270
Leu Pro Tyr Thr Pro Gly Val Pro Leu Val Ser Met Tyr Pro His Ala
        275                 280                 285
His Pro Leu Ala Ile Asp Leu Leu Gln Lys Met Leu Ile Phe Asp Pro
    290                 295                 300
Thr Lys Arg Ile Ser Val Thr Glu Ala Leu Glu His Pro Tyr Met Ser
305                 310                 315                 320
Pro Leu Tyr Asp Pro Ser Ala Asn Pro Pro Ala Gln Val Pro Ile Asp
                325                 330                 335
Leu Asp Ile Asp Glu Asn Ile Ser Ser Glu Met Ile Arg Glu Met Met
            340                 345                 350
Trp Gln Glu Met Leu His Tyr His Pro Glu Val Ala Thr Ala Ile Ser
        355                 360                 365
Met Ser
    370
```

<210> 23

<211> 20
<212> DNA
<213> Zea mays

<400> 23
atggcgatga tggtggatcc 20

<210> 24
<211> 20
<212> DNA
<213> Zea mays

<400> 24
tcatgacatg cttattgctg 20

<210> 25
<211> 1125
<212> DNA
<213> Zea mays

<220>
<221> CDS
<222> (1)...(1122)

<400> 25

```
atg agc gga gga ggc gtg gac gga gct ccg gtc gcc gag ttc cgg cag      48
Met Ser Gly Gly Gly Val Asp Gly Ala Pro Val Ala Glu Phe Arg Gln
 1               5                  10                  15

acg gtg acg cac ggc ggc cgg ttc ctg cag tac aac atc ttc ggc aac      96
Thr Val Thr His Gly Gly Arg Phe Leu Gln Tyr Asn Ile Phe Gly Asn
            20                  25                  30

ctg ttc gag atc acg cac aag tac cag ccc ccc atc atg ccc atc ggc     144
Leu Phe Glu Ile Thr His Lys Tyr Gln Pro Pro Ile Met Pro Ile Gly
        35                  40                  45

cgc ggc gcc tac ggg atc gtc tgc tcg gtg atg aac tcc gag acg aag     192
Arg Gly Ala Tyr Gly Ile Val Cys Ser Val Met Asn Ser Glu Thr Lys
    50                  55                  60

gag atg gtg gcc atc aag aag atc gcc aac gcc ttc gac aac cac atg     240
Glu Met Val Ala Ile Lys Lys Ile Ala Asn Ala Phe Asp Asn His Met
65                  70                  75                  80

gac gcc aag cgc acg ctc cgg gag atc aag ctg ctg agg cac ctc gac     288
Asp Ala Lys Arg Thr Leu Arg Glu Ile Lys Leu Leu Arg His Leu Asp
                85                  90                  95

cac gag aac atc atc ggc atc agg gac gtg atc ccg ccg ccc gtc ccg     336
His Glu Asn Ile Ile Gly Ile Arg Asp Val Ile Pro Pro Pro Val Pro
            100                 105                 110

caa gcg ttc aac gac gtg tac atc ggg acg gag ctg atg gac acg gac     384
Gln Ala Phe Asn Asp Val Tyr Ile Gly Thr Glu Leu Met Asp Thr Asp
        115                 120                 125

ctg cac cac atc atc cgg tcc aac cag gag ctc tcg gag gag cac tcc     432
Leu His His Ile Ile Arg Ser Asn Gln Glu Leu Ser Glu Glu His Ser
    130                 135                 140

cag tac ttc atg tac cag atc ctc cgc ggg ctc aag tac atc cac tcc     480
```

```
Gln Tyr Phe Met Tyr Gln Ile Leu Arg Gly Leu Lys Tyr Ile His Ser
145                 150                 155                 160

gcc aac gtg atc cac cgc gac ctc aag ccc agc aac ctg ctg gtg aac      528
Ala Asn Val Ile His Arg Asp Leu Lys Pro Ser Asn Leu Leu Val Asn
                165                 170                 175

gcc aac tgc gac ctc aag atc tgc gac ttc ggg ctg gcg cgc ccg tcc      576
Ala Asn Cys Asp Leu Lys Ile Cys Asp Phe Gly Leu Ala Arg Pro Ser
            180                 185                 190

tcc gag agc gac atg atg acg gag tac gtg gtc acg cgc tgg tac cgc      624
Ser Glu Ser Asp Met Met Thr Glu Tyr Val Val Thr Arg Trp Tyr Arg
        195                 200                 205

gcg ccc gag ctg ctg ctc aac tcc acc gac tac tcg gcg gcc atc gac      672
Ala Pro Glu Leu Leu Leu Asn Ser Thr Asp Tyr Ser Ala Ala Ile Asp
    210                 215                 220

gtc tgg tcc gtc ggc tgc atc ttc atg gag ctc atc aac cgc cag ccg      720
Val Trp Ser Val Gly Cys Ile Phe Met Glu Leu Ile Asn Arg Gln Pro
225                 230                 235                 240

ctc ttc ccc gga cgc gac cac atg cac cag atg cgc ctc ata acc gag      768
Leu Phe Pro Gly Arg Asp His Met His Gln Met Arg Leu Ile Thr Glu
                245                 250                 255

gtg atc ggg acg ccc acg gac gat gag ctc ggg ttc atc cgg aac gag      816
Val Ile Gly Thr Pro Thr Asp Asp Glu Leu Gly Phe Ile Arg Asn Glu
            260                 265                 270

gac gcg cgg aag tac atg cgc cac ctc ccg cag ttc ccg cgc cgg ccg      864
Asp Ala Arg Lys Tyr Met Arg His Leu Pro Gln Phe Pro Arg Arg Pro
        275                 280                 285

ttc gtc agc ctg ttc ccg cgg atg cag ccc gtc gcg ctg gac ctc atc      912
Phe Val Ser Leu Phe Pro Arg Met Gln Pro Val Ala Leu Asp Leu Ile
    290                 295                 300

gag cgg atg ctc acc ttc aac ccg ctg cag agg atc aca gtg gaa gag      960
Glu Arg Met Leu Thr Phe Asn Pro Leu Gln Arg Ile Thr Val Glu Glu
305                 310                 315                 320

gcg ctg gag cac ccg tac ctg gaa cgg cta cac gac gtc gcc gac gag     1008
Ala Leu Glu His Pro Tyr Leu Glu Arg Leu His Asp Val Ala Asp Glu
                325                 330                 335

ccc atc tgc acg gac ccg ttc tcg ttc gac ttc gag cag cag gct ctg     1056
Pro Ile Cys Thr Asp Pro Phe Ser Phe Asp Phe Glu Gln Gln Ala Leu
            340                 345                 350

acg gaa gac caa atg aag cag ctg ata ttc aac gag gcc atg gaa ctc     1104
Thr Glu Asp Gln Met Lys Gln Leu Ile Phe Asn Glu Ala Met Glu Leu
        355                 360                 365

aac ccc aac ttc cga tac tag                                        1125
Asn Pro Asn Phe Arg Tyr
    370
```

<210> 26
<211> 374

<212> PRT
<213> Zea mays

<400> 26

```
Met Ser Gly Gly Gly Val Asp Gly Ala Pro Val Ala Glu Phe Arg Gln
 1               5                  10                  15
Thr Val Thr His Gly Gly Arg Phe Leu Gln Tyr Asn Ile Phe Gly Asn
            20                  25                  30
Leu Phe Glu Ile Thr His Lys Tyr Gln Pro Pro Ile Met Pro Ile Gly
        35                  40                  45
Arg Gly Ala Tyr Gly Ile Val Cys Ser Val Met Asn Ser Glu Thr Lys
    50                  55                  60
Glu Met Val Ala Ile Lys Lys Ile Ala Asn Ala Phe Asp Asn His Met
65                  70                  75                  80
Asp Ala Lys Arg Thr Leu Arg Glu Ile Lys Leu Leu Arg His Leu Asp
                85                  90                  95
His Glu Asn Ile Ile Gly Ile Arg Asp Val Ile Pro Pro Pro Val Pro
            100                 105                 110
Gln Ala Phe Asn Asp Val Tyr Ile Gly Thr Glu Leu Met Asp Thr Asp
        115                 120                 125
Leu His His Ile Ile Arg Ser Asn Gln Glu Leu Ser Glu Glu His Ser
    130                 135                 140
Gln Tyr Phe Met Tyr Gln Ile Leu Arg Gly Leu Lys Tyr Ile His Ser
145                 150                 155                 160
Ala Asn Val Ile His Arg Asp Leu Lys Pro Ser Asn Leu Leu Val Asn
                165                 170                 175
Ala Asn Cys Asp Leu Lys Ile Cys Asp Phe Gly Leu Ala Arg Pro Ser
            180                 185                 190
Ser Glu Ser Asp Met Met Thr Glu Tyr Val Val Thr Arg Trp Tyr Arg
        195                 200                 205
Ala Pro Glu Leu Leu Leu Asn Ser Thr Asp Tyr Ser Ala Ala Ile Asp
    210                 215                 220
Val Trp Ser Val Gly Cys Ile Phe Met Glu Leu Ile Asn Arg Gln Pro
225                 230                 235                 240
Leu Phe Pro Gly Arg Asp His Met His Gln Met Arg Leu Ile Thr Glu
                245                 250                 255
Val Ile Gly Thr Pro Thr Asp Asp Glu Leu Gly Phe Ile Arg Asn Glu
            260                 265                 270
Asp Ala Arg Lys Tyr Met Arg His Leu Pro Gln Phe Pro Arg Arg Pro
        275                 280                 285
Phe Val Ser Leu Phe Pro Arg Met Gln Pro Val Ala Leu Asp Leu Ile
    290                 295                 300
Glu Arg Met Leu Thr Phe Asn Pro Leu Gln Arg Ile Thr Val Glu Glu
305                 310                 315                 320
Ala Leu Glu His Pro Tyr Leu Glu Arg Leu His Asp Val Ala Asp Glu
                325                 330                 335
Pro Ile Cys Thr Asp Pro Phe Ser Phe Asp Phe Glu Gln Gln Ala Leu
            340                 345                 350
Thr Glu Asp Gln Met Lys Gln Leu Ile Phe Asn Glu Ala Met Glu Leu
        355                 360                 365
Asn Pro Asn Phe Arg Tyr
    370
```

<210> 27
<211> 20
<212> DNA
<213> Zea mays

<400> 27
atgagcggag gaggcgtgga 20

<210> 28

<211> 20
<212> DNA
<213> Zea mays

<400> 28
ctagtatcgg aagttggggt 20

<210> 29
<211> 1197
<212> DNA
<213> Zea mays

<220>
<221> CDS
<222> (1)...(1194)

<400> 29

```
atg gac ggc ggg ggg cag ccc ccg gac acc gag atg tcg gag gcc ggc      48
Met Asp Gly Gly Gly Gln Pro Pro Asp Thr Glu Met Ser Glu Ala Gly
 1               5                  10                  15

gcg ggc ggc gga ggg cag ccg ccg cag cag cca ctg ccg ccg gtg ggc      96
Ala Gly Gly Gly Gly Gln Pro Pro Gln Gln Pro Leu Pro Pro Val Gly
            20                  25                  30

ggg ggc gtg atg ttg gac aac atc cag gcg acg ctc agc cac ggc ggc     144
Gly Gly Val Met Leu Asp Asn Ile Gln Ala Thr Leu Ser His Gly Gly
        35                  40                  45

cgc ttc atc cag tac aac atc ttc ggc aac gtg ttc gag gtc acc gcc     192
Arg Phe Ile Gln Tyr Asn Ile Phe Gly Asn Val Phe Glu Val Thr Ala
    50                  55                  60

aag tac aag ccc ccc gtc ctc ccc atc ggc aag ggc gcc tac ggc atc     240
Lys Tyr Lys Pro Pro Val Leu Pro Ile Gly Lys Gly Ala Tyr Gly Ile
65                  70                  75                  80

gtc tgc tcg gcg ctc aac tcc gag acg gcg gag cag gtg gcc atc aag     288
Val Cys Ser Ala Leu Asn Ser Glu Thr Ala Glu Gln Val Ala Ile Lys
                85                  90                  95

aag atc gcc aac gcc ttc gac aac aag atc gat gcc aag cgc acg ctc     336
Lys Ile Ala Asn Ala Phe Asp Asn Lys Ile Asp Ala Lys Arg Thr Leu
            100                 105                 110

cgc gag atc aag ctg ctc cgc cac atg gac cac gag aat att gtt gca     384
Arg Glu Ile Lys Leu Leu Arg His Met Asp His Glu Asn Ile Val Ala
        115                 120                 125

ata agg gga atc ata cct cct gcg cag agg gct gca ttc aat gat gtg     432
Ile Arg Gly Ile Ile Pro Pro Ala Gln Arg Ala Ala Phe Asn Asp Val
    130                 135                 140

tat att gca tat gaa ttg atg gat act gat ctg cat caa att att cgt     480
Tyr Ile Ala Tyr Glu Leu Met Asp Thr Asp Leu His Gln Ile Ile Arg
145                 150                 155                 160

tca aat caa gct ttg tca gag gag cac tgt cag tat ttt ctt tat caa     528
Ser Asn Gln Ala Leu Ser Glu Glu His Cys Gln Tyr Phe Leu Tyr Gln
                165                 170                 175

att ctt cgt ggc ttg aag tat ata cat tca gca aat gtt ctt cac cgt     576
```

```
Ile Leu Arg Gly Leu Lys Tyr Ile His Ser Ala Asn Val Leu His Arg
            180                 185                 190

gac ttg aag cct agc aat ctt ctt ttg aat gca aac tgt gac ctc aag      624
Asp Leu Lys Pro Ser Asn Leu Leu Leu Asn Ala Asn Cys Asp Leu Lys
        195                 200                 205

ata tgt gat ttt ggg ctt gct cgc acc acc tca gaa act gat ttt atg      672
Ile Cys Asp Phe Gly Leu Ala Arg Thr Thr Ser Glu Thr Asp Phe Met
    210                 215                 220

act gaa tat gtt gtc aca aga tgg tat aga gca cca gag ctt ttg ttg      720
Thr Glu Tyr Val Val Thr Arg Trp Tyr Arg Ala Pro Glu Leu Leu Leu
225                 230                 235                 240

aac tcc tca gaa tat act gct gcc atc gat gtg tgg tct gtg ggc tgt      768
Asn Ser Ser Glu Tyr Thr Ala Ala Ile Asp Val Trp Ser Val Gly Cys
                245                 250                 255

ata ttt atg gaa ctg atg gac cgc aaa ccc ttg ttt cct gga aga gac      816
Ile Phe Met Glu Leu Met Asp Arg Lys Pro Leu Phe Pro Gly Arg Asp
            260                 265                 270

cat gtc cat cag cta cgt ctt cta atg gag ctc att ggc acg cca aat      864
His Val His Gln Leu Arg Leu Leu Met Glu Leu Ile Gly Thr Pro Asn
        275                 280                 285

gag ggt gat ctt gat ttt gta aat gaa aat gca aga aga tat atc cgc      912
Glu Gly Asp Leu Asp Phe Val Asn Glu Asn Ala Arg Arg Tyr Ile Arg
    290                 295                 300

caa ctt cct cgt cac cct aga cag tcc tta cct gaa aaa ttt cca cat      960
Gln Leu Pro Arg His Pro Arg Gln Ser Leu Pro Glu Lys Phe Pro His
305                 310                 315                 320

gta caa ccc tta gca att gac ctg gtg gaa aag atg ctg act ttt gat     1008
Val Gln Pro Leu Ala Ile Asp Leu Val Glu Lys Met Leu Thr Phe Asp
                325                 330                 335

cct aga cag aga ata act gtt gaa ggc gca ctt gca cac cct tac ttg     1056
Pro Arg Gln Arg Ile Thr Val Glu Gly Ala Leu Ala His Pro Tyr Leu
            340                 345                 350

gca tca ctt cat gac ata agt gat gag cca ggc tgc tca atg ccc ttc     1104
Ala Ser Leu His Asp Ile Ser Asp Glu Pro Gly Cys Ser Met Pro Phe
        355                 360                 365

agc ttt gac ttc gag cag cat gca ttg tct gaa gaa caa atg aag gat     1152
Ser Phe Asp Phe Glu Gln His Ala Leu Ser Glu Glu Gln Met Lys Asp
    370                 375                 380

ctt atc tac caa gag gct ctt gca ttc aac cct gat tat cag             1194
Leu Ile Tyr Gln Glu Ala Leu Ala Phe Asn Pro Asp Tyr Gln
385                 390                 395

tag                                                                 1197
```

<210> 30
<211> 398
<212> PRT
<213> Zea mays

<400> 30

```
Met Asp Gly Gly Gly Gln Pro Pro Asp Thr Glu Met Ser Glu Ala Gly
 1               5                  10                  15
Ala Gly Gly Gly Gly Gln Pro Pro Gln Gln Pro Leu Pro Pro Val Gly
            20                  25                  30
Gly Gly Val Met Leu Asp Asn Ile Gln Ala Thr Leu Ser His Gly Gly
        35                  40                  45
Arg Phe Ile Gln Tyr Asn Ile Phe Gly Asn Val Phe Glu Val Thr Ala
    50                  55                  60
Lys Tyr Lys Pro Pro Val Leu Pro Ile Gly Lys Gly Ala Tyr Gly Ile
65                  70                  75                  80
Val Cys Ser Ala Leu Asn Ser Glu Thr Ala Glu Gln Val Ala Ile Lys
                85                  90                  95
Lys Ile Ala Asn Ala Phe Asp Asn Lys Ile Asp Ala Lys Arg Thr Leu
            100                 105                 110
Arg Glu Ile Lys Leu Leu Arg His Met Asp His Glu Asn Ile Val Ala
        115                 120                 125
Ile Arg Gly Ile Ile Pro Pro Ala Gln Arg Ala Ala Phe Asn Asp Val
    130                 135                 140
Tyr Ile Ala Tyr Glu Leu Met Asp Thr Asp Leu His Gln Ile Ile Arg
145                 150                 155                 160
Ser Asn Gln Ala Leu Ser Glu Glu His Cys Gln Tyr Phe Leu Tyr Gln
                165                 170                 175
Ile Leu Arg Gly Leu Lys Tyr Ile His Ser Ala Asn Val Leu His Arg
            180                 185                 190
Asp Leu Lys Pro Ser Asn Leu Leu Leu Asn Ala Asn Cys Asp Leu Lys
        195                 200                 205
Ile Cys Asp Phe Gly Leu Ala Arg Thr Thr Ser Glu Thr Asp Phe Met
    210                 215                 220
Thr Glu Tyr Val Val Thr Arg Trp Tyr Arg Ala Pro Glu Leu Leu Leu
225                 230                 235                 240
Asn Ser Ser Glu Tyr Thr Ala Ala Ile Asp Val Trp Ser Val Gly Cys
                245                 250                 255
Ile Phe Met Glu Leu Met Asp Arg Lys Pro Leu Phe Pro Gly Arg Asp
            260                 265                 270
His Val His Gln Leu Arg Leu Leu Met Glu Leu Ile Gly Thr Pro Asn
        275                 280                 285
Glu Gly Asp Leu Asp Phe Val Asn Glu Asn Ala Arg Arg Tyr Ile Arg
    290                 295                 300
Gln Leu Pro Arg His Pro Arg Gln Ser Leu Pro Glu Lys Phe Pro His
305                 310                 315                 320
Val Gln Pro Leu Ala Ile Asp Leu Val Glu Lys Met Leu Thr Phe Asp
                325                 330                 335
Pro Arg Gln Arg Ile Thr Val Glu Gly Ala Leu Ala His Pro Tyr Leu
            340                 345                 350
Ala Ser Leu His Asp Ile Ser Asp Glu Pro Gly Cys Ser Met Pro Phe
        355                 360                 365
Ser Phe Asp Phe Glu Gln His Ala Leu Ser Glu Glu Gln Met Lys Asp
    370                 375                 380
Leu Ile Tyr Gln Glu Ala Leu Ala Phe Asn Pro Asp Tyr Gln
385                 390                 395
```

<210> 31
<211> 20
<212> DNA
<213> Zea mays

<400> 31
atggacggcg gggggcagcc 20

<210> 32
<211> 20

<212> DNA
<213> Zea mays

<400> 32
ctactgataa tcagggttga 20

<210> 33
<211> 1197
<212> DNA
<213> Zea mays

<220>
<221> CDS
<222> (1)...(1194)

<400> 33

```
atg gac ggc ggg ggg cag ccc ccg gac acg gag atg aca gac gcc ggc  48
Met Asp Gly Gly Gly Gln Pro Pro Asp Thr Glu Met Thr Asp Ala Gly
 1               5                  10                  15

ttg ggc ggc ggg ggg caa ccg ccg ccg ccg ccg caa cag ccg gcg ggc  96
Leu Gly Gly Gly Gly Gln Pro Pro Pro Pro Pro Gln Gln Pro Ala Gly
            20                  25                  30

ggg gcc ggg atg atg gag aac atc cac gcg acg ctc agc cac ggt ggc 144
Gly Ala Gly Met Met Glu Asn Ile His Ala Thr Leu Ser His Gly Gly
        35                  40                  45

cgc ttc atc cag tac aac atc ttc ggc aac gtg ttc gag gtc acc tcc 192
Arg Phe Ile Gln Tyr Asn Ile Phe Gly Asn Val Phe Glu Val Thr Ser
    50                  55                  60

aag tac aag ccc ccc atc ctc ccc atc ggc aag ggc gcc tac ggc atc 240
Lys Tyr Lys Pro Pro Ile Leu Pro Ile Gly Lys Gly Ala Tyr Gly Ile
65                  70                  75                  80

gtc tgc tcg gcg ctc aac tcc gag acg gca gag cag gtg gcc atc aag 288
Val Cys Ser Ala Leu Asn Ser Glu Thr Ala Glu Gln Val Ala Ile Lys
                85                  90                  95

aag atc gcc aac gcc ttc gac aac aag atc gat gcc aag cgc acg ctc 336
Lys Ile Ala Asn Ala Phe Asp Asn Lys Ile Asp Ala Lys Arg Thr Leu
            100                 105                 110

cgc gag atc aag ctg ctc cgc cac atg gac cac gag aat att gtt gca 384
Arg Glu Ile Lys Leu Leu Arg His Met Asp His Glu Asn Ile Val Ala
        115                 120                 125

ata aga gat atc ata cct cct cca ttg agg gag gca ttc aat gat gtg 432
Ile Arg Asp Ile Ile Pro Pro Pro Leu Arg Glu Ala Phe Asn Asp Val
    130                 135                 140

tat att gcc tat gaa ttg atg gat act gat ctg cat caa att att cgt 480
Tyr Ile Ala Tyr Glu Leu Met Asp Thr Asp Leu His Gln Ile Ile Arg
145                 150                 155                 160

tca aat caa gct ttg tca gag gag cac tgt cag tat ttt ctt tat caa 528
Ser Asn Gln Ala Leu Ser Glu Glu His Cys Gln Tyr Phe Leu Tyr Gln
                165                 170                 175

att ctt cgt ggc ttg aag tat ata cat tca gca aat gtc ctt cac cgt 576
Ile Leu Arg Gly Leu Lys Tyr Ile His Ser Ala Asn Val Leu His Arg
```

```
            180                 185                 190

gac ttg aag cct agc aat ctt ctt ttg aat gca aac tgt gac ctc aag      624
Asp Leu Lys Pro Ser Asn Leu Leu Leu Asn Ala Asn Cys Asp Leu Lys
        195                 200                 205

ata tgt gat ttt ggg ctt gct cgc acc acc tca gaa act gat ttt atg      672
Ile Cys Asp Phe Gly Leu Ala Arg Thr Thr Ser Glu Thr Asp Phe Met
    210                 215                 220

act gaa tat gtt gtc aca aga tgg tat aga gca cca gag ctt tta ttg      720
Thr Glu Tyr Val Val Thr Arg Trp Tyr Arg Ala Pro Glu Leu Leu Leu
225                 230                 235                 240

aac tcc tct gaa tat act gct gcc att gat gtg tgg tct gtg ggc tgt      768
Asn Ser Ser Glu Tyr Thr Ala Ala Ile Asp Val Trp Ser Val Gly Cys
                245                 250                 255

ata ttt atg gaa ctg atg gac cga aaa ccc ttg ttt cct gga aga gat      816
Ile Phe Met Glu Leu Met Asp Arg Lys Pro Leu Phe Pro Gly Arg Asp
            260                 265                 270

cat gtc cat cag cta cgt cta cta atg gag ctc att gga aca ccg aat      864
His Val His Gln Leu Arg Leu Leu Met Glu Leu Ile Gly Thr Pro Asn
        275                 280                 285

gag gct gat ctt gat ttt gta aat gaa aat gca aga aga tat atc cgc      912
Glu Ala Asp Leu Asp Phe Val Asn Glu Asn Ala Arg Arg Tyr Ile Arg
    290                 295                 300

caa ctt ccc tgt cat gct aga cag tcc ttc cct gaa aaa ttt cca cat      960
Gln Leu Pro Cys His Ala Arg Gln Ser Phe Pro Glu Lys Phe Pro His
305                 310                 315                 320

gta caa cct tta gca att gac cta gtg gaa aag atg cta act ttt gat     1008
Val Gln Pro Leu Ala Ile Asp Leu Val Glu Lys Met Leu Thr Phe Asp
                325                 330                 335

cct aga cag aga ata act gtt gaa ggc gca ctt gca cac cct tac ttg     1056
Pro Arg Gln Arg Ile Thr Val Glu Gly Ala Leu Ala His Pro Tyr Leu
            340                 345                 350

gca tca ctt cat gac ata agt gat gag cca gtc tgc tca atg ccc ttc     1104
Ala Ser Leu His Asp Ile Ser Asp Glu Pro Val Cys Ser Met Pro Phe
        355                 360                 365

agc ttc gac ttc gag cag cat gca tta tct gaa gaa cag atg aag gat     1152
Ser Phe Asp Phe Glu Gln His Ala Leu Ser Glu Glu Gln Met Lys Asp
    370                 375                 380

ctg atc tac caa gag gct ctt gca ttc aac cca gat tac cag             1194
Leu Ile Tyr Gln Glu Ala Leu Ala Phe Asn Pro Asp Tyr Gln
385                 390                 395

tag                                                                 1197
```

<210> 34
<211> 398
<212> PRT
<213> Zea mays

<400> 34

```
Met Asp Gly Gly Gly Gln Pro Pro Asp Thr Glu Met Thr Asp Ala Gly
 1               5                  10                  15
Leu Gly Gly Gly Gly Gln Pro Pro Pro Pro Pro Gln Gln Pro Ala Gly
            20                  25                  30
Gly Ala Gly Met Met Glu Asn Ile His Ala Thr Leu Ser His Gly Gly
        35                  40                  45
Arg Phe Ile Gln Tyr Asn Ile Phe Gly Asn Val Phe Glu Val Thr Ser
    50                  55                  60
Lys Tyr Lys Pro Pro Ile Leu Pro Ile Gly Lys Gly Ala Tyr Gly Ile
65                  70                  75                  80
Val Cys Ser Ala Leu Asn Ser Glu Thr Ala Glu Gln Val Ala Ile Lys
                85                  90                  95
Lys Ile Ala Asn Ala Phe Asp Asn Lys Ile Asp Ala Lys Arg Thr Leu
            100                 105                 110
Arg Glu Ile Lys Leu Leu Arg His Met Asp His Glu Asn Ile Val Ala
        115                 120                 125
Ile Arg Asp Ile Ile Pro Pro Pro Leu Arg Glu Ala Phe Asn Asp Val
    130                 135                 140
Tyr Ile Ala Tyr Glu Leu Met Asp Thr Asp Leu His Gln Ile Ile Arg
145                 150                 155                 160
Ser Asn Gln Ala Leu Ser Glu Glu His Cys Gln Tyr Phe Leu Tyr Gln
                165                 170                 175
Ile Leu Arg Gly Leu Lys Tyr Ile His Ser Ala Asn Val Leu His Arg
            180                 185                 190
Asp Leu Lys Pro Ser Asn Leu Leu Leu Asn Ala Asn Cys Asp Leu Lys
        195                 200                 205
Ile Cys Asp Phe Gly Leu Ala Arg Thr Thr Ser Glu Thr Asp Phe Met
    210                 215                 220
Thr Glu Tyr Val Val Thr Arg Trp Tyr Arg Ala Pro Glu Leu Leu Leu
225                 230                 235                 240
Asn Ser Ser Glu Tyr Thr Ala Ala Ile Asp Val Trp Ser Val Gly Cys
                245                 250                 255
Ile Phe Met Glu Leu Met Asp Arg Lys Pro Leu Phe Pro Gly Arg Asp
            260                 265                 270
His Val His Gln Leu Arg Leu Leu Met Glu Leu Ile Gly Thr Pro Asn
        275                 280                 285
Glu Ala Asp Leu Asp Phe Val Asn Glu Asn Ala Arg Arg Tyr Ile Arg
    290                 295                 300
Gln Leu Pro Cys His Ala Arg Gln Ser Phe Pro Glu Lys Phe Pro His
305                 310                 315                 320
Val Gln Pro Leu Ala Ile Asp Leu Val Glu Lys Met Leu Thr Phe Asp
                325                 330                 335
Pro Arg Gln Arg Ile Thr Val Glu Gly Ala Leu Ala His Pro Tyr Leu
            340                 345                 350
Ala Ser Leu His Asp Ile Ser Asp Glu Pro Val Cys Ser Met Pro Phe
        355                 360                 365
Ser Phe Asp Phe Glu Gln His Ala Leu Ser Glu Glu Gln Met Lys Asp
    370                 375                 380
Leu Ile Tyr Gln Glu Ala Leu Ala Phe Asn Pro Asp Tyr Gln
385                 390                 395
```

<210> 35
<211> 20
<212> DNA
<213> Zea mays

<400> 35
atggacggcg gggggcagcc 20

<210> 36
<211> 20
<212> DNA

<213> Zea mays

<400> 36
ctactggtaa tctgggttga 20

<210> 37
<211> 1053
<212> DNA
<213> Zea mays

<220>
<221> CDS
<222> (1)...(1050)

<400> 37

```
atg gcg acg cca cgg aag ccg atc aag ctc acg ctg ccg tcc cac gag        48
Met Ala Thr Pro Arg Lys Pro Ile Lys Leu Thr Leu Pro Ser His Glu
 1               5                  10                  15

acc acc atc ggc aag ttc ctg acg cac agc ggg acg ttc acg gac ggg        96
Thr Thr Ile Gly Lys Phe Leu Thr His Ser Gly Thr Phe Thr Asp Gly
            20                  25                  30

gat ctg cgc gtc aac aag gac ggc ctc cgc atc gtc tcg cgg agg gag       144
Asp Leu Arg Val Asn Lys Asp Gly Leu Arg Ile Val Ser Arg Arg Glu
        35                  40                  45

gga ggc gag gct cct cct ata gag ccg ttg gat agt caa ctg agc tta       192
Gly Gly Glu Ala Pro Pro Ile Glu Pro Leu Asp Ser Gln Leu Ser Leu
    50                  55                  60

gat gat cta gac gtt ata aaa gtg atc ggg aaa ggt agc agc gga aat       240
Asp Asp Leu Asp Val Ile Lys Val Ile Gly Lys Gly Ser Ser Gly Asn
65                  70                  75                  80

gtg caa ttg gtc cgc cac aaa ttt act ggc cag ttt ttt gct ctg aag       288
Val Gln Leu Val Arg His Lys Phe Thr Gly Gln Phe Phe Ala Leu Lys
                85                  90                  95

gtt att caa cta aat att gat gag agt ata cgc aaa cag att gcc aag       336
Val Ile Gln Leu Asn Ile Asp Glu Ser Ile Arg Lys Gln Ile Ala Lys
            100                 105                 110

gag ttg aag ata aac tta tca aca cag tgc caa tat gtt gtt gtg ttc       384
Glu Leu Lys Ile Asn Leu Ser Thr Gln Cys Gln Tyr Val Val Val Phe
        115                 120                 125

tat cag tgt ttc tat ttc aat ggt gcc att tct att gtt ttg gaa tac       432
Tyr Gln Cys Phe Tyr Phe Asn Gly Ala Ile Ser Ile Val Leu Glu Tyr
    130                 135                 140

atg gat ggt ggc tcc ctt gca gat ttc ctg aag act gtt aaa acc att       480
Met Asp Gly Gly Ser Leu Ala Asp Phe Leu Lys Thr Val Lys Thr Ile
145                 150                 155                 160

cca gag gcc tac ctc gct gct atc tgt acg cag atg cta aaa gga ctg       528
Pro Glu Ala Tyr Leu Ala Ala Ile Cys Thr Gln Met Leu Lys Gly Leu
                165                 170                 175

atc tat ttg cat aac gag aag cgc gtt ata cac cga gat ctg aaa cca       576
Ile Tyr Leu His Asn Glu Lys Arg Val Ile His Arg Asp Leu Lys Pro
            180                 185                 190
```

```
tca aat ata ttg ata aat cat agg ggt gaa gta aaa ata tca gat ttt      624
Ser Asn Ile Leu Ile Asn His Arg Gly Glu Val Lys Ile Ser Asp Phe
        195                 200                 205

ggt gtg agt gcc att ata tct agt tcc tct tcg caa cga gat aca ttt      672
Gly Val Ser Ala Ile Ile Ser Ser Ser Ser Ser Gln Arg Asp Thr Phe
    210                 215                 220

att ggc aca cgc aac tac atg gcg cca gaa aga atc gat gga aag aaa      720
Ile Gly Thr Arg Asn Tyr Met Ala Pro Glu Arg Ile Asp Gly Lys Lys
225                 230                 235                 240

cat ggt tct atg agt gat atc tgg agt ttg gga cta gtg ata ctg gaa      768
His Gly Ser Met Ser Asp Ile Trp Ser Leu Gly Leu Val Ile Leu Glu
                245                 250                 255

tgt gca acc ggc atc ttt cca ttt cct cct tgt gaa agc ttc tac gaa      816
Cys Ala Thr Gly Ile Phe Pro Phe Pro Pro Cys Glu Ser Phe Tyr Glu
            260                 265                 270

ctt ctc gtg gct gtt gtt gat caa ccg cca cct tct gcg ccg ccg gat      864
Leu Leu Val Ala Val Val Asp Gln Pro Pro Pro Ser Ala Pro Pro Asp
        275                 280                 285

cag ttt tca cca gaa ttc tgt ggg ttc att tct gca tgt ctc cag aag      912
Gln Phe Ser Pro Glu Phe Cys Gly Phe Ile Ser Ala Cys Leu Gln Lys
    290                 295                 300

gat gct aat gac agg tca tca gcc caa gcc tta ttg gac cat ccg ttc      960
Asp Ala Asn Asp Arg Ser Ser Ala Gln Ala Leu Leu Asp His Pro Phe
305                 310                 315                 320

ctg agc atg tat gat gac ctg cat gta gat ctt gct tcg tac ttc acg     1008
Leu Ser Met Tyr Asp Asp Leu His Val Asp Leu Ala Ser Tyr Phe Thr
                325                 330                 335

aca gca gga tct cct ctc gcc acc ttc aat tcc agg caa ctc              1050
Thr Ala Gly Ser Pro Leu Ala Thr Phe Asn Ser Arg Gln Leu
            340                 345                 350

taa                                                                 1053
```

<210> 38
<211> 350
<212> PRT
<213> Zea mays

<400> 38

```
Met Ala Thr Pro Arg Lys Pro Ile Lys Leu Thr Leu Pro Ser His Glu
 1               5                  10                  15
Thr Thr Ile Gly Lys Phe Leu Thr His Ser Gly Thr Phe Thr Asp Gly
            20                  25                  30
Asp Leu Arg Val Asn Lys Asp Gly Leu Arg Ile Val Ser Arg Arg Glu
        35                  40                  45
Gly Gly Glu Ala Pro Pro Ile Glu Pro Leu Asp Ser Gln Leu Ser Leu
    50                  55                  60
Asp Asp Leu Asp Val Ile Lys Val Ile Gly Lys Gly Ser Ser Gly Asn
65                  70                  75                  80
Val Gln Leu Val Arg His Lys Phe Thr Gly Gln Phe Phe Ala Leu Lys
                85                  90                  95
Val Ile Gln Leu Asn Ile Asp Glu Ser Ile Arg Lys Gln Ile Ala Lys
```

```
            100                 105                 110
Glu Leu Lys Ile Asn Leu Ser Thr Gln Cys Gln Tyr Val Val Val Phe
        115                 120                 125
Tyr Gln Cys Phe Tyr Phe Asn Gly Ala Ile Ser Ile Val Leu Glu Tyr
    130                 135                 140
Met Asp Gly Gly Ser Leu Ala Asp Phe Leu Lys Thr Val Lys Thr Ile
145                 150                 155                 160
Pro Glu Ala Tyr Leu Ala Ala Ile Cys Thr Gln Met Leu Lys Gly Leu
                165                 170                 175
Ile Tyr Leu His Asn Glu Lys Arg Val Ile His Arg Asp Leu Lys Pro
            180                 185                 190
Ser Asn Ile Leu Ile Asn His Arg Gly Glu Val Lys Ile Ser Asp Phe
        195                 200                 205
Gly Val Ser Ala Ile Ile Ser Ser Ser Ser Ser Gln Arg Asp Thr Phe
    210                 215                 220
Ile Gly Thr Arg Asn Tyr Met Ala Pro Glu Arg Ile Asp Gly Lys Lys
225                 230                 235                 240
His Gly Ser Met Ser Asp Ile Trp Ser Leu Gly Leu Val Ile Leu Glu
                245                 250                 255
Cys Ala Thr Gly Ile Phe Pro Phe Pro Pro Cys Glu Ser Phe Tyr Glu
            260                 265                 270
Leu Leu Val Ala Val Val Asp Gln Pro Pro Pro Ser Ala Pro Pro Asp
        275                 280                 285
Gln Phe Ser Pro Glu Phe Cys Gly Phe Ile Ser Ala Cys Leu Gln Lys
    290                 295                 300
Asp Ala Asn Asp Arg Ser Ser Ala Gln Ala Leu Leu Asp His Pro Phe
305                 310                 315                 320
Leu Ser Met Tyr Asp Asp Leu His Val Asp Leu Ala Ser Tyr Phe Thr
                325                 330                 335
Thr Ala Gly Ser Pro Leu Ala Thr Phe Asn Ser Arg Gln Leu
            340                 345                 350
```

<210> 39
<211> 20
<212> DNA
<213> Zea mays

<400> 39
atggcgacgc cacggaagcc 20

<210> 40
<211> 20
<212> DNA
<213> Zea mays

<400> 40
ttagagttgc ctggaattga 20

<210> 41
<211> 846
<212> DNA
<213> Zea mays

<220>
<221> CDS
<222> (1)...(843)

<400> 41
```
atg gct caa gag ctc cac gaa acg tcc tct tgc tct gcc acc acc acc      48
Met Ala Gln Glu Leu His Glu Thr Ser Ser Cys Ser Ala Thr Thr Thr
 1               5                  10                  15
```

```
tcg tcg tgc acc aca tcc tgc tgc tcg tcc act gtc aca gac tcg tcc        96
Ser Ser Cys Thr Thr Ser Cys Cys Ser Ser Thr Val Thr Asp Ser Ser
                20                  25                  30

tct tcg ccc ccg tca ccg gcg gcg gcc aat gcc gcg ccc gcg aca cgg       144
Ser Ser Pro Pro Ser Pro Ala Ala Ala Asn Ala Ala Pro Ala Thr Arg
            35                  40                  45

aag cgg cag gcg ttg gag gcc gag gcc gag gcc gag gcg ggc ggt gag       192
Lys Arg Gln Ala Leu Glu Ala Glu Ala Glu Ala Glu Ala Gly Gly Glu
        50                  55                  60

gag gag gag gag gag gag gga agc tgt gct ggt aat aaa gcg gcg ccg       240
Glu Glu Glu Glu Glu Glu Gly Ser Cys Ala Gly Asn Lys Ala Ala Pro
 65                  70                  75                  80

gcc aag aag cga ccg cgg ggc agc gag ggg aaa cac ccg acg ttc cgc       288
Ala Lys Lys Arg Pro Arg Gly Ser Glu Gly Lys His Pro Thr Phe Arg
                85                  90                  95

ggc gtg cgg atg cgg gcg tgg ggc aag tgg gtg tcg gag atc cgc gag       336
Gly Val Arg Met Arg Ala Trp Gly Lys Trp Val Ser Glu Ile Arg Glu
            100                 105                 110

ccg cgc aag aag tcg cgc ata tgg ctc ggc acg ttc ccc acc gcc gag       384
Pro Arg Lys Lys Ser Arg Ile Trp Leu Gly Thr Phe Pro Thr Ala Glu
        115                 120                 125

atg gcc gcg cgc gcc cac gac gtc gcg gcg ctc gcc atc aaa ggc cgc       432
Met Ala Ala Arg Ala His Asp Val Ala Ala Leu Ala Ile Lys Gly Arg
    130                 135                 140

gcc gcg cac ctc aac ttc ccg gac ttc gcc ggc gcg ctc ccg cgc gcc       480
Ala Ala His Leu Asn Phe Pro Asp Phe Ala Gly Ala Leu Pro Arg Ala
145                 150                 155                 160

gcg tcc gcg gcg ccc aag gac gtc cag gca gcc gcc gca ttg gcc gct       528
Ala Ser Ala Ala Pro Lys Asp Val Gln Ala Ala Ala Ala Leu Ala Ala
                165                 170                 175

gcg ttc acg tcg ccg tca tcg gag ccc ggc gcc ggc gcg cac gag gag       576
Ala Phe Thr Ser Pro Ser Ser Glu Pro Gly Ala Gly Ala His Glu Glu
            180                 185                 190

ccc gct gcc aag gac ggc gcc gcg ccc gag gag gca gcc gcc gac gca       624
Pro Ala Ala Lys Asp Gly Ala Ala Pro Glu Glu Ala Ala Ala Asp Ala
        195                 200                 205

cag gca cca gta cca gta gca cta cca ccg ccg gcg gcc tct cgg cca       672
Gln Ala Pro Val Pro Val Ala Leu Pro Pro Pro Ala Ala Ser Arg Pro
    210                 215                 220

ggg acg ccg tcg agc ggc gtg gag gac gag cgg cag ctg ttc gac ctg       720
Gly Thr Pro Ser Ser Gly Val Glu Asp Glu Arg Gln Leu Phe Asp Leu
225                 230                 235                 240

ccg gac ctg ctc ctc gac atc cgg gac ggg ttc ggg cgc ttc ccg ccg       768
Pro Asp Leu Leu Leu Asp Ile Arg Asp Gly Phe Gly Arg Phe Pro Pro
                245                 250                 255

atg tgg gcc ccg ctc act gac gtg gag gag gtg gtc aat gcg gag ctg       816
Met Trp Ala Pro Leu Thr Asp Val Glu Glu Val Val Asn Ala Glu Leu
```

260 265 270

cgc ctc gag gag ccg ctg ctt tgg gag tag 846
Arg Leu Glu Glu Pro Leu Leu Trp Glu
275 280

<210> 42
<211> 281
<212> PRT
<213> Zea mays

<400> 42
Met Ala Gln Glu Leu His Glu Thr Ser Ser Cys Ser Ala Thr Thr Thr
1 5 10 15
Ser Ser Cys Thr Thr Ser Cys Cys Ser Ser Thr Val Thr Asp Ser Ser
20 25 30
Ser Ser Pro Pro Ser Pro Ala Ala Ala Asn Ala Ala Pro Ala Thr Arg
35 40 45
Lys Arg Gln Ala Leu Glu Ala Glu Ala Glu Ala Glu Ala Gly Gly Glu
50 55 60
Glu Glu Glu Glu Glu Glu Gly Ser Cys Ala Gly Asn Lys Ala Ala Pro
65 70 75 80
Ala Lys Lys Arg Pro Arg Gly Ser Glu Gly Lys His Pro Thr Phe Arg
85 90 95
Gly Val Arg Met Arg Ala Trp Gly Lys Trp Val Ser Glu Ile Arg Glu
100 105 110
Pro Arg Lys Lys Ser Arg Ile Trp Leu Gly Thr Phe Pro Thr Ala Glu
115 120 125
Met Ala Ala Arg Ala His Asp Val Ala Ala Leu Ala Ile Lys Gly Arg
130 135 140
Ala Ala His Leu Asn Phe Pro Asp Phe Ala Gly Ala Leu Pro Arg Ala
145 150 155 160
Ala Ser Ala Ala Pro Lys Asp Val Gln Ala Ala Ala Ala Leu Ala Ala
165 170 175
Ala Phe Thr Ser Pro Ser Ser Glu Pro Gly Ala Gly Ala His Glu Glu
180 185 190
Pro Ala Ala Lys Asp Gly Ala Ala Pro Glu Glu Ala Ala Ala Asp Ala
195 200 205
Gln Ala Pro Val Pro Val Ala Leu Pro Pro Pro Ala Ala Ser Arg Pro
210 215 220
Gly Thr Pro Ser Ser Gly Val Glu Asp Glu Arg Gln Leu Phe Asp Leu
225 230 235 240
Pro Asp Leu Leu Leu Asp Ile Arg Asp Gly Phe Gly Arg Phe Pro Pro
245 250 255
Met Trp Ala Pro Leu Thr Asp Val Glu Glu Val Val Asn Ala Glu Leu
260 265 270
Arg Leu Glu Glu Pro Leu Leu Trp Glu
275 280

<210> 43
<211> 20
<212> DNA
<213> Zea mays

<400> 43
atggctcaag agctccacga 20

<210> 44
<211> 20
<212> DNA
<213> Zea mays

<400> 44
ctactcccaa agcagcggct 20

<210> 45
<211> 2257
<212> DNA
<213> Zea mays

<220>
<221> CDS
<222> (108)...(1844)

<400> 45
gtcgacccac gcgtccgtgc ttggagcggg tggacgtgtg gtgtgtatat agttaagcgg 60
gcgggcgggc gggcagccaa tgccgtcggg cgcgcagggc taggagc atg ggc gga 116
Met Gly Gly
1

ggg ggc gcg atc gtg cac ggc ttc cgc cgc tgg ttc cac cgc cgg agc 164
Gly Gly Ala Ile Val His Gly Phe Arg Arg Trp Phe His Arg Arg Ser
5 10 15

gaa ggc tcc acc tcc aac tcc aac cag tcg tcc ctc gcc ggc gag ggg 212
Glu Gly Ser Thr Ser Asn Ser Asn Gln Ser Ser Leu Ala Gly Glu Gly
20 25 30 35

gaa gcg gac gac ggc tca tcc gat ctc gag gtc atc gag gac ccg gat 260
Glu Ala Asp Asp Gly Ser Ser Asp Leu Glu Val Ile Glu Asp Pro Asp
40 45 50

ctc ctc ggc ctc cgc gcc atc cgc gtg ccc aag cgc aag atg ccg ctg 308
Leu Leu Gly Leu Arg Ala Ile Arg Val Pro Lys Arg Lys Met Pro Leu
55 60 65

ccc gtc gag agc cac agg aag aac tca gta gaa atg gag ttc ttc aca 356
Pro Val Glu Ser His Arg Lys Asn Ser Val Glu Met Glu Phe Phe Thr
70 75 80

gag tac gga gag gca agc cag tac caa atc caa gaa gtc att ggt aaa 404
Glu Tyr Gly Glu Ala Ser Gln Tyr Gln Ile Gln Glu Val Ile Gly Lys
85 90 95

ggg agc tac gga gta gtt gct gca gct gtg gac acc cgc acc ggt gag 452
Gly Ser Tyr Gly Val Val Ala Ala Ala Val Asp Thr Arg Thr Gly Glu
100 105 110 115

cgg gtt gct atc aag aag atc aat gat gtg ttt gag cac gtc tcg gat 500
Arg Val Ala Ile Lys Lys Ile Asn Asp Val Phe Glu His Val Ser Asp
120 125 130

gcc aca cgc ata ctg cgg gag atc aag ctc ctc cgg ttg ctt cgt cac 548
Ala Thr Arg Ile Leu Arg Glu Ile Lys Leu Leu Arg Leu Leu Arg His
135 140 145

cca gac ata gtc gag att aag cac atc atg ctc cct cct tct cgg agg 596
Pro Asp Ile Val Glu Ile Lys His Ile Met Leu Pro Pro Ser Arg Arg
150 155 160

gag ttc caa gac atc tat gtt gtt ttt gag ctc atg gag tcg gat ctc 644
Glu Phe Gln Asp Ile Tyr Val Val Phe Glu Leu Met Glu Ser Asp Leu
165 170 175

```
cat caa gtg atc agg gcg aat gat gac ctc acc cct gaa cat cac cag      692
His Gln Val Ile Arg Ala Asn Asp Asp Leu Thr Pro Glu His His Gln
180                 185                 190                 195

ttt ttc ctg tac caa ctt ctt cga gct ctc aag tac ata cat gcc gct      740
Phe Phe Leu Tyr Gln Leu Leu Arg Ala Leu Lys Tyr Ile His Ala Ala
                200                 205                 210

aat gta ttc cat cgt gat cta aag ccc aag aat ata ctg gct aac tcg      788
Asn Val Phe His Arg Asp Leu Lys Pro Lys Asn Ile Leu Ala Asn Ser
            215                 220                 225

gac tgc aaa ctg aaa ata tgt gac ttt gga ctt gcc cgt gca tca ttc      836
Asp Cys Lys Leu Lys Ile Cys Asp Phe Gly Leu Ala Arg Ala Ser Phe
        230                 235                 240

aat gat gcc cct tca gct ata ttt tgg aca gac tat gtg gca aca agg      884
Asn Asp Ala Pro Ser Ala Ile Phe Trp Thr Asp Tyr Val Ala Thr Arg
    245                 250                 255

tgg tac cgg gca cct gaa cta tgt ggt tcc ttt ttc tcc aaa tac act      932
Trp Tyr Arg Ala Pro Glu Leu Cys Gly Ser Phe Phe Ser Lys Tyr Thr
260                 265                 270                 275

cct gca att gat att tgg agc att ggg tgc ata ttt gct gaa ctt ctc      980
Pro Ala Ile Asp Ile Trp Ser Ile Gly Cys Ile Phe Ala Glu Leu Leu
                280                 285                 290

act gga agg cca cta ttt cct ggg aag aat gtt gta cac caa tta gat     1028
Thr Gly Arg Pro Leu Phe Pro Gly Lys Asn Val Val His Gln Leu Asp
            295                 300                 305

att ata aca gat ctt ctt ggg act cca tca tct gag acc tta tcc cgg     1076
Ile Ile Thr Asp Leu Leu Gly Thr Pro Ser Ser Glu Thr Leu Ser Arg
        310                 315                 320

att cga aat gag aag gcc agg aga tat ttg agc tgc atg cgg aaa aaa     1124
Ile Arg Asn Glu Lys Ala Arg Arg Tyr Leu Ser Cys Met Arg Lys Lys
    325                 330                 335

tat ccc gtg ccc ttt act cat aaa ttt cga cat gct gat cct ttg gct     1172
Tyr Pro Val Pro Phe Thr His Lys Phe Arg His Ala Asp Pro Leu Ala
340                 345                 350                 355

ctt cgt ctt cta gag cgt ttg ctt gca ttt gat cct aaa gac cgg cct     1220
Leu Arg Leu Leu Glu Arg Leu Leu Ala Phe Asp Pro Lys Asp Arg Pro
                360                 365                 370

agt gca gaa gag gct cta gct gat cca tac ttt gca tcc ctt gct aat     1268
Ser Ala Glu Glu Ala Leu Ala Asp Pro Tyr Phe Ala Ser Leu Ala Asn
            375                 380                 385

gtg gaa cgt gag cca tca aga cac ccg att tct aaa ctt gag ttt gaa     1316
Val Glu Arg Glu Pro Ser Arg His Pro Ile Ser Lys Leu Glu Phe Glu
        390                 395                 400

ttt gaa aga aag aag ctg gcc aaa gat gat gtt aga gaa ctg atc tat     1364
Phe Glu Arg Lys Lys Leu Ala Lys Asp Asp Val Arg Glu Leu Ile Tyr
    405                 410                 415

cga gag att ttg gag tat cat cca cag atg ctg gag gag tac atg aaa     1412
```

```
Arg Glu Ile Leu Glu Tyr His Pro Gln Met Leu Glu Glu Tyr Met Lys
420                 425                 430                 435

ggc gga gaa cag att agc ttc ctc tat cca agc gga gtc gac cgc ttc     1460
Gly Gly Glu Gln Ile Ser Phe Leu Tyr Pro Ser Gly Val Asp Arg Phe
                440                 445                 450

aaa ctg cag ttc gct cat ctt gag gag cat tac agc aaa gga gaa aga     1508
Lys Leu Gln Phe Ala His Leu Glu Glu His Tyr Ser Lys Gly Glu Arg
            455                 460                 465

ggt tcc cca ctg caa aga aag cat gct tcc tta cca agg gaa aga gtg     1556
Gly Ser Pro Leu Gln Arg Lys His Ala Ser Leu Pro Arg Glu Arg Val
        470                 475                 480

gtt gta tcc aag gat ggc aat act gaa caa caa att aat gac cag gag     1604
Val Val Ser Lys Asp Gly Asn Thr Glu Gln Gln Ile Asn Asp Gln Glu
    485                 490                 495

aga agc gca gat tct gtt gcc cgc act acc gtg agc cct cca agg tca     1652
Arg Ser Ala Asp Ser Val Ala Arg Thr Thr Val Ser Pro Pro Arg Ser
500                 505                 510                 515

gag gat gct gac atg gat gat gtg aaa ccc aca agc tta agc tct cgg     1700
Glu Asp Ala Asp Met Asp Asp Val Lys Pro Thr Ser Leu Ser Ser Arg
                520                 525                 530

agc tac ctg aag agt gca agc ata agc gct tcc aag tgc gtc gtt gtc     1748
Ser Tyr Leu Lys Ser Ala Ser Ile Ser Ala Ser Lys Cys Val Val Val
            535                 540                 545

aga aat aaa cac cca gag cca gag gat gat gga ata cct gag gag atg     1796
Arg Asn Lys His Pro Glu Pro Glu Asp Asp Gly Ile Pro Glu Glu Met
        550                 555                 560

gaa ggc acc gtc gat gga ttg tct gaa aag gtc tcc agg atg cac tcc     1844
Glu Gly Thr Val Asp Gly Leu Ser Glu Lys Val Ser Arg Met His Ser
    565                 570                 575

taaggcagga ctgtccatgg cgacagcaat ctcgacacaa gatttctgtc tggtgcactg   1904
taatgccgac gacacactgt ttggatcttg tgtacaccca cagctagttc aactaagatg   1964
tagcactagt tcgtccatgg aggaagcaga atggacgatg aagtgtggtc agcctgtcga   2024
ccatgaggtg ttttcttcca atggagtcta gggtatgttt gagttctaca acggcttgtg   2084
ccgtacgaag tgtcagcagt gtacgaggag gaggatatcg ttgtcagtac acattttgcc   2144
tgtctagaaa attatgagta ctttctccaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa   2204
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaagggcggc cgc          2257

      <210> 46
      <211> 579
      <212> PRT
      <213> Zea mays

      <400> 46
Met Gly Gly Gly Gly Ala Ile Val His Gly Phe Arg Arg Trp Phe His
 1               5                  10                  15
Arg Arg Ser Glu Gly Ser Thr Ser Asn Ser Asn Gln Ser Ser Leu Ala
            20                  25                  30
Gly Glu Gly Glu Ala Asp Asp Gly Ser Ser Asp Leu Glu Val Ile Glu
        35                  40                  45
Asp Pro Asp Leu Leu Gly Leu Arg Ala Ile Arg Val Pro Lys Arg Lys
    50                  55                  60
Met Pro Leu Pro Val Glu Ser His Arg Lys Asn Ser Val Glu Met Glu
```

```
65                  70                  75                  80
Phe Phe Thr Glu Tyr Gly Glu Ala Ser Gln Tyr Gln Ile Gln Glu Val
                85                  90                  95
Ile Gly Lys Gly Ser Tyr Gly Val Val Ala Ala Ala Val Asp Thr Arg
            100                 105                 110
Thr Gly Glu Arg Val Ala Ile Lys Lys Ile Asn Asp Val Phe Glu His
        115                 120                 125
Val Ser Asp Ala Thr Arg Ile Leu Arg Glu Ile Lys Leu Leu Arg Leu
    130                 135                 140
Leu Arg His Pro Asp Ile Val Glu Ile Lys His Ile Met Leu Pro Pro
145                 150                 155                 160
Ser Arg Arg Glu Phe Gln Asp Ile Tyr Val Val Phe Glu Leu Met Glu
                165                 170                 175
Ser Asp Leu His Gln Val Ile Arg Ala Asn Asp Asp Leu Thr Pro Glu
            180                 185                 190
His His Gln Phe Phe Leu Tyr Gln Leu Leu Arg Ala Leu Lys Tyr Ile
        195                 200                 205
His Ala Ala Asn Val Phe His Arg Asp Leu Lys Pro Lys Asn Ile Leu
    210                 215                 220
Ala Asn Ser Asp Cys Lys Leu Lys Ile Cys Asp Phe Gly Leu Ala Arg
225                 230                 235                 240
Ala Ser Phe Asn Asp Ala Pro Ser Ala Ile Phe Trp Thr Asp Tyr Val
                245                 250                 255
Ala Thr Arg Trp Tyr Arg Ala Pro Glu Leu Cys Gly Ser Phe Phe Ser
            260                 265                 270
Lys Tyr Thr Pro Ala Ile Asp Ile Trp Ser Ile Gly Cys Ile Phe Ala
        275                 280                 285
Glu Leu Leu Thr Gly Arg Pro Leu Phe Pro Gly Lys Asn Val Val His
    290                 295                 300
Gln Leu Asp Ile Ile Thr Asp Leu Leu Gly Thr Pro Ser Ser Glu Thr
305                 310                 315                 320
Leu Ser Arg Ile Arg Asn Glu Lys Ala Arg Arg Tyr Leu Ser Cys Met
                325                 330                 335
Arg Lys Lys Tyr Pro Val Pro Phe Thr His Lys Phe Arg His Ala Asp
            340                 345                 350
Pro Leu Ala Leu Arg Leu Leu Glu Arg Leu Leu Ala Phe Asp Pro Lys
        355                 360                 365
Asp Arg Pro Ser Ala Glu Glu Ala Leu Ala Asp Pro Tyr Phe Ala Ser
    370                 375                 380
Leu Ala Asn Val Glu Arg Glu Pro Ser Arg His Pro Ile Ser Lys Leu
385                 390                 395                 400
Glu Phe Glu Phe Glu Arg Lys Lys Leu Ala Lys Asp Asp Val Arg Glu
                405                 410                 415
Leu Ile Tyr Arg Glu Ile Leu Glu Tyr His Pro Gln Met Leu Glu Glu
            420                 425                 430
Tyr Met Lys Gly Gly Glu Gln Ile Ser Phe Leu Tyr Pro Ser Gly Val
        435                 440                 445
Asp Arg Phe Lys Leu Gln Phe Ala His Leu Glu Glu His Tyr Ser Lys
    450                 455                 460
Gly Glu Arg Gly Ser Pro Leu Gln Arg Lys His Ala Ser Leu Pro Arg
465                 470                 475                 480
Glu Arg Val Val Val Ser Lys Asp Gly Asn Thr Glu Gln Gln Ile Asn
                485                 490                 495
Asp Gln Glu Arg Ser Ala Asp Ser Val Ala Arg Thr Thr Val Ser Pro
            500                 505                 510
Pro Arg Ser Glu Asp Ala Asp Met Asp Asp Val Lys Pro Thr Ser Leu
        515                 520                 525
Ser Ser Arg Ser Tyr Leu Lys Ser Ala Ser Ile Ser Ala Ser Lys Cys
    530                 535                 540
Val Val Val Arg Asn Lys His Pro Glu Pro Glu Asp Asp Gly Ile Pro
545                 550                 555                 560
Glu Glu Met Glu Gly Thr Val Asp Gly Leu Ser Glu Lys Val Ser Arg
```

565 570 575
Met His Ser

<210> 47
<211> 20
<212> DNA
<213> Zea mays

<400> 47
atgggcggag ggggcgcgat 20

<210> 48
<211> 20
<212> DNA
<213> Zea mays

<400> 48
ttaggagtgc atcctggaga 20

<210> 49
<211> 938
<212> DNA
<213> Zea mays

<220>
<221> CDS
<222> (39)...(743)

<400> 49
gaattcggca cgagcagaaa gcgaagcacc agcctctg atg cca agc caa agc aca 56
Met Pro Ser Gln Ser Thr
1 5

cac acc tcc tcc agc ctt tct tcc tct aca acc cca att ctt ccc tgc 104
His Thr Ser Ser Ser Leu Ser Ser Ser Thr Thr Pro Ile Leu Pro Cys
10 15 20

agc tct gct gct gcc ttg ttt cga tcg gtc atg gcg gcg gtg gcc aca 152
Ser Ser Ala Ala Ala Leu Phe Arg Ser Val Met Ala Ala Val Ala Thr
25 30 35

gag act ccg ttc cat gtc ctg gcg gtg gac gac agc ctc ccg gac agg 200
Glu Thr Pro Phe His Val Leu Ala Val Asp Asp Ser Leu Pro Asp Arg
40 45 50

aag ctc atc gag agg ctc ctc aag acc tct tcc ttc caa gtg acc act 248
Lys Leu Ile Glu Arg Leu Leu Lys Thr Ser Ser Phe Gln Val Thr Thr
55 60 65 70

gtc gac tcc ggg agc aag gcg ctg cag ttc ctg ggc ctc cat gac cag 296
Val Asp Ser Gly Ser Lys Ala Leu Gln Phe Leu Gly Leu His Asp Gln
75 80 85

gac agc acg gtt cct cct gtc cac acg cac cag ctg gat gtg gct gcc 344
Asp Ser Thr Val Pro Pro Val His Thr His Gln Leu Asp Val Ala Ala
90 95 100

aat cag gat gtg gct gtg aac ctg atc atc aca gac tac tgc atg cct 392
Asn Gln Asp Val Ala Val Asn Leu Ile Ile Thr Asp Tyr Cys Met Pro
105 110 115

```
ggc atg aca gga tat gac ctg ctc aag aag atc aag gag tcg tcg tct      440
Gly Met Thr Gly Tyr Asp Leu Leu Lys Lys Ile Lys Glu Ser Ser Ser
    120                 125                 130

ctc aga gat atc ccg gtg gtg atc atg tcc tct gag aac att cct tca      488
Leu Arg Asp Ile Pro Val Val Ile Met Ser Ser Glu Asn Ile Pro Ser
135                 140                 145                 150

agg atc aat agg tgc ctg gag gaa gga gct gac gag ttc ttc cta aaa      536
Arg Ile Asn Arg Cys Leu Glu Glu Gly Ala Asp Glu Phe Phe Leu Lys
                155                 160                 165

cct gtg cgg cta tca gac atg aac aag ctc aag ccc cac ata ctg aaa      584
Pro Val Arg Leu Ser Asp Met Asn Lys Leu Lys Pro His Ile Leu Lys
            170                 175                 180

agc aga tgc aac cag gaa cag cac cag caa agt gac agt cac agt ggc      632
Ser Arg Cys Asn Gln Glu Gln His Gln Gln Ser Asp Ser His Ser Gly
        185                 190                 195

gaa cgc agg aac ccc aca atc agc agc agc gat agc ata aac aac cgc      680
Glu Arg Arg Asn Pro Thr Ile Ser Ser Ser Asp Ser Ile Asn Asn Arg
    200                 205                 210

aag aga aag ggt gca ggc acc gaa gaa atc ttg ccc cag ctg gca aac      728
Lys Arg Lys Gly Ala Gly Thr Glu Glu Ile Leu Pro Gln Leu Ala Asn
215                 220                 225                 230

aga tca agg cac agt taactgagaa ctgactagta cagctaaaac ctttttctttt      783
Arg Ser Arg His Ser
                235

tcactttact ttggcttact gatttgtaac gtagatgtta gccctggatt tgcaaaacgg     843
aacggaatct gtaaactgat actgcttgag tcgaatcgat cgaagtgatt ctgcaccaga     903
tctccctaaa aaaaaaaaaa aaaaaaaaac tcgag                                938

      <210> 50
      <211> 235
      <212> PRT
      <213> Zea mays

      <400> 50
Met Pro Ser Gln Ser Thr His Thr Ser Ser Ser Leu Ser Ser Ser Thr
 1               5                  10                  15
Thr Pro Ile Leu Pro Cys Ser Ser Ala Ala Ala Leu Phe Arg Ser Val
            20                  25                  30
Met Ala Ala Val Ala Thr Glu Thr Pro Phe His Val Leu Ala Val Asp
        35                  40                  45
Asp Ser Leu Pro Asp Arg Lys Leu Ile Glu Arg Leu Leu Lys Thr Ser
    50                  55                  60
Ser Phe Gln Val Thr Thr Val Asp Ser Gly Ser Lys Ala Leu Gln Phe
65                  70                  75                  80
Leu Gly Leu His Asp Gln Asp Ser Thr Val Pro Pro Val His Thr His
                85                  90                  95
Gln Leu Asp Val Ala Ala Asn Gln Asp Val Ala Val Asn Leu Ile Ile
            100                 105                 110
Thr Asp Tyr Cys Met Pro Gly Met Thr Gly Tyr Asp Leu Leu Lys Lys
        115                 120                 125
Ile Lys Glu Ser Ser Ser Leu Arg Asp Ile Pro Val Val Ile Met Ser
    130                 135                 140
Ser Glu Asn Ile Pro Ser Arg Ile Asn Arg Cys Leu Glu Glu Gly Ala
145                 150                 155                 160
```

```
Asp Glu Phe Phe Leu Lys Pro Val Arg Leu Ser Asp Met Asn Lys Leu
                  165                 170                 175
Lys Pro His Ile Leu Lys Ser Arg Cys Asn Gln Glu Gln His Gln Gln
             180                 185                 190
Ser Asp Ser His Ser Gly Glu Arg Arg Asn Pro Thr Ile Ser Ser Ser
         195                 200                 205
Asp Ser Ile Asn Asn Arg Lys Arg Lys Gly Ala Gly Thr Glu Glu Ile
     210                 215                 220
Leu Pro Gln Leu Ala Asn Arg Ser Arg His Ser
225                 230                 235
```

<210> 51
<211> 20
<212> DNA
<213> Zea mays

<400> 51
atgccaagcc aaagcacaca 20

<210> 52
<211> 20
<212> DNA
<213> Zea mays

<400> 52
ttaactgtgc cttgatctgt 20

<210> 53
<211> 1322
<212> DNA
<213> Zea mays

<220>
<221> CDS
<222> (217)...(864)

<400> 53

```
gtcgacccac gcgtccggta aaaagccacc taccctaccc tcctgcgctg gtactactta     60
ctacctagcc aggccaagga gcccaacgag ggggaacgct gcggtaggcg cgcgctgctc    120
gtcccacccc acgccgccat gacggtgcca gatgccgagt cgcgcttcca tgtcctcgcg    180
gtggacgaca gcctcgtcga caggaaactc atcgag atg ctg ctc aag acc tcg      234
                                      Met Leu Leu Lys Thr Ser
                                       1               5

tcc tac caa gtg acc acg gtg gat tcc ggg agc aag gcg ctg gag ctg      282
Ser Tyr Gln Val Thr Thr Val Asp Ser Gly Ser Lys Ala Leu Glu Leu
            10                  15                  20

ctg ggg ctg agg gac gcg tcg tcg ccg tct ccg tcc tcg cct gac cac      330
Leu Gly Leu Arg Asp Ala Ser Ser Pro Ser Pro Ser Ser Pro Asp His
        25                  30                  35

cag gag atc gac gtg aat ctc atc atc act gac tac tgc atg cca ggc      378
Gln Glu Ile Asp Val Asn Leu Ile Ile Thr Asp Tyr Cys Met Pro Gly
    40                  45                  50

atg aca gga tac gat ctg ctc aag aga atg aag ggg tcc tct tcg ctc      426
Met Thr Gly Tyr Asp Leu Leu Lys Arg Met Lys Gly Ser Ser Ser Leu
55                  60                  65                  70

aag gac att cct gtg gtg atc atg tcg tct gag aat gtg cct gcc cgg      474
Lys Asp Ile Pro Val Val Ile Met Ser Ser Glu Asn Val Pro Ala Arg
```

```
                    75                  80                  85

atc agc agg tgc ttg caa gac ggc gcg gag gag ttc ttc ctg aag ccc      522
Ile Ser Arg Cys Leu Gln Asp Gly Ala Glu Glu Phe Phe Leu Lys Pro
            90                  95                  100

gtg aag ctg gcc gac atg aag aag ctc aag tcg cac ctg ctg aaa cgg      570
Val Lys Leu Ala Asp Met Lys Lys Leu Lys Ser His Leu Leu Lys Arg
        105                 110                 115

aag cag ccc aag gag gcg cag gcg cag gcg cag gcg cag gcg cag cag      618
Lys Gln Pro Lys Glu Ala Gln Ala Gln Ala Gln Ala Gln Ala Gln Gln
    120                 125                 130

gga cag gcg gtg gag ctg gag cct gag cag cag ctg gac ccg cgc acg      666
Gly Gln Ala Val Glu Leu Glu Pro Glu Gln Gln Leu Asp Pro Arg Thr
135                 140                 145                 150

cag ccg gcg cac gac gcg gag gaa acc gcg gca gag ccg ccg ccg gcc      714
Gln Pro Ala His Asp Ala Glu Glu Thr Ala Ala Glu Pro Pro Pro Ala
                155                 160                 165

gca tcc aac gga acc acc gat ggc ggc aac aag agg aag gcg gca gcc      762
Ala Ser Asn Gly Thr Thr Asp Gly Gly Asn Lys Arg Lys Ala Ala Ala
            170                 175                 180

atg gag gag gag ggg atg ctg gcc gtg atg acg gtg gcg gcg ccg gag      810
Met Glu Glu Glu Gly Met Leu Ala Val Met Thr Val Ala Ala Pro Glu
        185                 190                 195

agc agc acc aag ccg agg ctg tcc acc acc acc agc agc ctg gcg gtg      858
Ser Ser Thr Lys Pro Arg Leu Ser Thr Thr Thr Ser Ser Leu Ala Val
    200                 205                 210

gaa acc tgagctgaga atcggacggc gaccggatca catattacta cctacgtaca       914
Glu Thr
215

ttaatccacc attgcgggca gtcgagaggc aaccaaccaa ccaagaacca accgaccaaa    974
cccaggccac cgccgagaaa gagctcgtcc tccttttctt tcgctggttc ttttggttct   1034
cggagctaat taatggaatg aaggtgtgag ctcctgggga tggaggggcg gttacgggct   1094
ctaagtgaag aacaattacg acgtaccttа cctagcggct gctgcttcgg ccggcaaatt   1154
ttgttgtgct cgtggtcatg gacacccttg tgttcgttag gagggggtcg gggtggaacg   1214
gtgaaaagga agctgtagat agaaaaggca ggcgtctcat gctcttgttg tctagtctaa   1274
tttttaattt tactcggctt attaaaaaaa aaaaaaaagg gcggccgc               1322

      <210> 54
      <211> 216
      <212> PRT
      <213> Zea mays

      <400> 54
Met Leu Leu Lys Thr Ser Ser Tyr Gln Val Thr Thr Val Asp Ser Gly
 1               5                  10                  15
Ser Lys Ala Leu Glu Leu Leu Gly Leu Arg Asp Ala Ser Ser Pro Ser
            20                  25                  30
Pro Ser Ser Pro Asp His Gln Glu Ile Asp Val Asn Leu Ile Ile Thr
        35                  40                  45
Asp Tyr Cys Met Pro Gly Met Thr Gly Tyr Asp Leu Leu Lys Arg Met
    50                  55                  60
Lys Gly Ser Ser Ser Leu Lys Asp Ile Pro Val Val Ile Met Ser Ser
65                  70                  75                  80
```

```
Glu Asn Val Pro Ala Arg Ile Ser Arg Cys Leu Gln Asp Gly Ala Glu
                85                  90                  95
Glu Phe Phe Leu Lys Pro Val Lys Leu Ala Asp Met Lys Lys Leu Lys
            100                 105                 110
Ser His Leu Leu Lys Arg Lys Gln Pro Lys Glu Ala Gln Ala Gln Ala
        115                 120                 125
Gln Ala Gln Ala Gln Gln Gly Gln Ala Val Glu Leu Glu Pro Glu Gln
    130                 135                 140
Gln Leu Asp Pro Arg Thr Gln Pro Ala His Asp Ala Glu Glu Thr Ala
145                 150                 155                 160
Ala Glu Pro Pro Pro Ala Ala Ser Asn Gly Thr Thr Asp Gly Gly Asn
                165                 170                 175
Lys Arg Lys Ala Ala Ala Met Glu Glu Glu Gly Met Leu Ala Val Met
            180                 185                 190
Thr Val Ala Ala Pro Glu Ser Ser Thr Lys Pro Arg Leu Ser Thr Thr
        195                 200                 205
Thr Ser Ser Leu Ala Val Glu Thr
    210                 215
```

<210> 55
<211> 20
<212> DNA
<213> Zea mays

<400> 55
atgctgctca agacctcgtc 20

<210> 56
<211> 20
<212> DNA
<213> Zea mays

<400> 56
tcaggtttcc accgccaggc 20

<210> 57
<211> 1635
<212> DNA
<213> Zea mays

<220>
<221> CDS
<222> (1)...(1635)

<400> 57

```
atg gat ttt ttc aca gag tac ggc gag ggg agc agg tac aag ata gaa      48
Met Asp Phe Phe Thr Glu Tyr Gly Glu Gly Ser Arg Tyr Lys Ile Glu
 1               5                  10                  15

gag gtc ata ggg aaa ggc agt tac ggt gtg gtt tgc tct gcc ttg gac      96
Glu Val Ile Gly Lys Gly Ser Tyr Gly Val Val Cys Ser Ala Leu Asp
            20                  25                  30

act cat act ggt gaa aag gtc gcc ata aag aag ata aat gac atc ttt     144
Thr His Thr Gly Glu Lys Val Ala Ile Lys Lys Ile Asn Asp Ile Phe
        35                  40                  45

gag cat gtg tct gat gcg aca agg ata ctt cgg gaa atc aag ttg ctt     192
Glu His Val Ser Asp Ala Thr Arg Ile Leu Arg Glu Ile Lys Leu Leu
    50                  55                  60

aga ctc ctg cgg cat ccg gat atc gtg gaa ata aag cat att cta ctt     240
```

```
Arg Leu Leu Arg His Pro Asp Ile Val Glu Ile Lys His Ile Leu Leu
 65                  70                  75                  80

cct cca tca agg agg gaa ttc agg gat att tac gtt gtg ttt gaa ctt      288
Pro Pro Ser Arg Arg Glu Phe Arg Asp Ile Tyr Val Val Phe Glu Leu
                 85                  90                  95

atg gag tct gat ttg cac caa gtt ata aag gct aat gat gac ttg acc      336
Met Glu Ser Asp Leu His Gln Val Ile Lys Ala Asn Asp Asp Leu Thr
            100                 105                 110

cca gag cat tat cag ttc ttc ttg tat cag ttg ctt cga ggt ttg aaa      384
Pro Glu His Tyr Gln Phe Phe Leu Tyr Gln Leu Leu Arg Gly Leu Lys
        115                 120                 125

tac ata cac aca gca aat gta ttc cat cga gat ctc aaa cca aaa aat      432
Tyr Ile His Thr Ala Asn Val Phe His Arg Asp Leu Lys Pro Lys Asn
    130                 135                 140

att ctg gca aat gct gat tgt aag ctc aaa ata tgt gat ttt ggc cta      480
Ile Leu Ala Asn Ala Asp Cys Lys Leu Lys Ile Cys Asp Phe Gly Leu
145                 150                 155                 160

gca aga gta gct ttc agt gat act cca acc gcg atc ttt tgg aca gat      528
Ala Arg Val Ala Phe Ser Asp Thr Pro Thr Ala Ile Phe Trp Thr Asp
                165                 170                 175

tat gtt gca act agg tgg tac cga gca ccg gag ctg tgt gga tca ttt      576
Tyr Val Ala Thr Arg Trp Tyr Arg Ala Pro Glu Leu Cys Gly Ser Phe
            180                 185                 190

ttc tca aag tat aca cca gcg ata gat ata tgg agt att ggt tgt ata      624
Phe Ser Lys Tyr Thr Pro Ala Ile Asp Ile Trp Ser Ile Gly Cys Ile
        195                 200                 205

ttt gcg gaa cta tta act gga aaa ccg ctt ttc ccg ggg aaa aat gtg      672
Phe Ala Glu Leu Leu Thr Gly Lys Pro Leu Phe Pro Gly Lys Asn Val
    210                 215                 220

gta cat caa ctt gat ata att aca gat ctc atg gga aca cct tct cca      720
Val His Gln Leu Asp Ile Ile Thr Asp Leu Met Gly Thr Pro Ser Pro
225                 230                 235                 240

gaa gca ata tcg agg att cgt aat gag aag gcc agg cgc tac ttg agt      768
Glu Ala Ile Ser Arg Ile Arg Asn Glu Lys Ala Arg Arg Tyr Leu Ser
                245                 250                 255

aat atg agg cgg aaa aag ccc ata cct ttt aca cag aag ttc cca aat      816
Asn Met Arg Arg Lys Lys Pro Ile Pro Phe Thr Gln Lys Phe Pro Asn
            260                 265                 270

gca gat cca ctt gca tta ggt ttg tta gag cga atg cta gct ttt gag      864
Ala Asp Pro Leu Ala Leu Gly Leu Leu Glu Arg Met Leu Ala Phe Glu
        275                 280                 285

cca aaa gac agg cca agt gct gaa gag gct ctt gct cat cct tat ttt      912
Pro Lys Asp Arg Pro Ser Ala Glu Glu Ala Leu Ala His Pro Tyr Phe
    290                 295                 300

aag aat ata gcc aat gtg gat aga gag cct tct gca caa gct gtc aca      960
Lys Asn Ile Ala Asn Val Asp Arg Glu Pro Ser Ala Gln Ala Val Thr
305                 310                 315                 320
```

```
aag ctc gaa ttt gag ttc gag aga cgg aga gtc act aaa gat gac ata      1008
Lys Leu Glu Phe Glu Phe Glu Arg Arg Arg Val Thr Lys Asp Asp Ile
                325                 330                 335

agg gaa ctt ata tat aga gaa att ctt gag tat cat cca aag atg ttg      1056
Arg Glu Leu Ile Tyr Arg Glu Ile Leu Glu Tyr His Pro Lys Met Leu
            340                 345                 350

aaa gaa ttt att gag gga act gag tca agt ggt ttc atg tac cca agt      1104
Lys Glu Phe Ile Glu Gly Thr Glu Ser Ser Gly Phe Met Tyr Pro Ser
        355                 360                 365

gca gta gat aat ttt aaa aag cag ttt gca tac ctt gag gaa cat tat      1152
Ala Val Asp Asn Phe Lys Lys Gln Phe Ala Tyr Leu Glu Glu His Tyr
    370                 375                 380

gcc aag gga tct acg gga act cca cct gag agg caa cat aac tca ttg      1200
Ala Lys Gly Ser Thr Gly Thr Pro Pro Glu Arg Gln His Asn Ser Leu
385                 390                 395                 400

cca cgg cct tct gtt gtc tat tca gac aac caa tca caa agc acg gcc      1248
Pro Arg Pro Ser Val Val Tyr Ser Asp Asn Gln Ser Gln Ser Thr Ala
                405                 410                 415

aat ata aca gag gat ctt tcc aag tgt ata atc aga gaa aat gca cag      1296
Asn Ile Thr Glu Asp Leu Ser Lys Cys Ile Ile Arg Glu Asn Ala Gln
            420                 425                 430

aag tca cag caa tat aat gct tca gtt gca aat aaa ttt cca tcc cat      1344
Lys Ser Gln Gln Tyr Asn Ala Ser Val Ala Asn Lys Phe Pro Ser His
        435                 440                 445

gtt cct caa ggt gct gct gca agg cct ggt aaa gca gtt ggt tcg gtg      1392
Val Pro Gln Gly Ala Ala Ala Arg Pro Gly Lys Ala Val Gly Ser Val
    450                 455                 460

ctg cag tac aac ctc tct caa gca tct gct gct gag caa tat gag cag      1440
Leu Gln Tyr Asn Leu Ser Gln Ala Ser Ala Ala Glu Gln Tyr Glu Gln
465                 470                 475                 480

cga agg gtt gct agg cac cca gca gtt gcc cca aac aac atc cct tct      1488
Arg Arg Val Ala Arg His Pro Ala Val Ala Pro Asn Asn Ile Pro Ser
                485                 490                 495

gga agc tca tac ccc aga aga aat cag acc tgc aag agc gag aca ggc      1536
Gly Ser Ser Tyr Pro Arg Arg Asn Gln Thr Cys Lys Ser Glu Thr Gly
            500                 505                 510

gac act gag agg atg gac gtg aac caa gct ggg cag ccg aag tcc tat      1584
Asp Thr Glu Arg Met Asp Val Asn Gln Ala Gly Gln Pro Lys Ser Tyr
        515                 520                 525

gca gca aac aaa cta cct gca act gtt gat ggt cgc ggc ggc cat tgg      1632
Ala Ala Asn Lys Leu Pro Ala Thr Val Asp Gly Arg Gly Gly His Trp
    530                 535                 540

tag                                                                 1635
*
```

<210> 58
<211> 544
<212> PRT
<213> Zea mays

<400> 58

```
Met Asp Phe Phe Thr Glu Tyr Gly Glu Gly Ser Arg Tyr Lys Ile Glu
 1               5                  10                  15
Glu Val Ile Gly Lys Gly Ser Tyr Gly Val Val Cys Ser Ala Leu Asp
            20                  25                  30
Thr His Thr Gly Glu Lys Val Ala Ile Lys Lys Ile Asn Asp Ile Phe
        35                  40                  45
Glu His Val Ser Asp Ala Thr Arg Ile Leu Arg Glu Ile Lys Leu Leu
    50                  55                  60
Arg Leu Leu Arg His Pro Asp Ile Val Glu Ile Lys His Ile Leu Leu
65                  70                  75                  80
Pro Pro Ser Arg Arg Glu Phe Arg Asp Ile Tyr Val Val Phe Glu Leu
                85                  90                  95
Met Glu Ser Asp Leu His Gln Val Ile Lys Ala Asn Asp Asp Leu Thr
            100                 105                 110
Pro Glu His Tyr Gln Phe Phe Leu Tyr Gln Leu Leu Arg Gly Leu Lys
        115                 120                 125
Tyr Ile His Thr Ala Asn Val Phe His Arg Asp Leu Lys Pro Lys Asn
    130                 135                 140
Ile Leu Ala Asn Ala Asp Cys Lys Leu Lys Ile Cys Asp Phe Gly Leu
145                 150                 155                 160
Ala Arg Val Ala Phe Ser Asp Thr Pro Thr Ala Ile Phe Trp Thr Asp
                165                 170                 175
Tyr Val Ala Thr Arg Trp Tyr Arg Ala Pro Glu Leu Cys Gly Ser Phe
            180                 185                 190
Phe Ser Lys Tyr Thr Pro Ala Ile Asp Ile Trp Ser Ile Gly Cys Ile
        195                 200                 205
Phe Ala Glu Leu Leu Thr Gly Lys Pro Leu Phe Pro Gly Lys Asn Val
    210                 215                 220
Val His Gln Leu Asp Ile Ile Thr Asp Leu Met Gly Thr Pro Ser Pro
225                 230                 235                 240
Glu Ala Ile Ser Arg Ile Arg Asn Glu Lys Ala Arg Arg Tyr Leu Ser
                245                 250                 255
Asn Met Arg Arg Lys Lys Pro Ile Pro Phe Thr Gln Lys Phe Pro Asn
            260                 265                 270
Ala Asp Pro Leu Ala Leu Gly Leu Leu Glu Arg Met Leu Ala Phe Glu
        275                 280                 285
Pro Lys Asp Arg Pro Ser Ala Glu Glu Ala Leu Ala His Pro Tyr Phe
    290                 295                 300
Lys Asn Ile Ala Asn Val Asp Arg Glu Pro Ser Ala Gln Ala Val Thr
305                 310                 315                 320
Lys Leu Glu Phe Glu Phe Glu Arg Arg Arg Val Thr Lys Asp Asp Ile
                325                 330                 335
Arg Glu Leu Ile Tyr Arg Glu Ile Leu Glu Tyr His Pro Lys Met Leu
            340                 345                 350
Lys Glu Phe Ile Glu Gly Thr Glu Ser Ser Gly Phe Met Tyr Pro Ser
        355                 360                 365
Ala Val Asp Asn Phe Lys Lys Gln Phe Ala Tyr Leu Glu Glu His Tyr
    370                 375                 380
Ala Lys Gly Ser Thr Gly Thr Pro Pro Glu Arg Gln His Asn Ser Leu
385                 390                 395                 400
Pro Arg Pro Ser Val Val Tyr Ser Asp Asn Gln Ser Gln Ser Thr Ala
                405                 410                 415
Asn Ile Thr Glu Asp Leu Ser Lys Cys Ile Ile Arg Glu Asn Ala Gln
            420                 425                 430
Lys Ser Gln Gln Tyr Asn Ala Ser Val Ala Asn Lys Phe Pro Ser His
        435                 440                 445
```

```
Val Pro Gln Gly Ala Ala Ala Arg Pro Gly Lys Ala Val Gly Ser Val
    450                 455                 460
Leu Gln Tyr Asn Leu Ser Gln Ala Ser Ala Ala Glu Gln Tyr Glu Gln
465                 470                 475                 480
Arg Arg Val Ala Arg His Pro Ala Val Ala Pro Asn Asn Ile Pro Ser
                485                 490                 495
Gly Ser Ser Tyr Pro Arg Arg Asn Gln Thr Cys Lys Ser Glu Thr Gly
            500                 505                 510
Asp Thr Glu Arg Met Asp Val Asn Gln Ala Gly Gln Pro Lys Ser Tyr
        515                 520                 525
Ala Ala Asn Lys Leu Pro Ala Thr Val Asp Gly Arg Gly Gly His Trp
    530                 535                 540
```

<210> 59
<211> 25
<212> DNA
<213> Zea mays

<400> 59
atggattttt tcacagagta cggcg 25

<210> 60
<211> 20
<212> DNA
<213> Zea mays

<400> 60
tcaatgcgaa accaaattta 20

<210> 61
<211> 4014
<212> DNA
<213> Zea mays

<220>
<221> CDS
<222> (1)...(4014)

<400> 61

```
atg gcg tcg agg cag cac aat gcg cag ttc cac aaa tcc aag acg ctt      48
Met Ala Ser Arg Gln His Asn Ala Gln Phe His Lys Ser Lys Thr Leu
 1               5                  10                  15

gac aac aaa tac atg ctt gga gat gaa ata ggg aag ggg gcg tac ggc      96
Asp Asn Lys Tyr Met Leu Gly Asp Glu Ile Gly Lys Gly Ala Tyr Gly
            20                  25                  30

cgc gta tac aag ggg ctt gac ctg gag aat ggc gat ttc gtg gcc atc     144
Arg Val Tyr Lys Gly Leu Asp Leu Glu Asn Gly Asp Phe Val Ala Ile
        35                  40                  45

aaa cag gtc tcg ctg gag aac att ccg cag gag gat ctc aac ata ata     192
Lys Gln Val Ser Leu Glu Asn Ile Pro Gln Glu Asp Leu Asn Ile Ile
    50                  55                  60

atg caa gag atc gac ctt ttg aaa aat ctt aat cat aaa aat att gtc     240
Met Gln Glu Ile Asp Leu Leu Lys Asn Leu Asn His Lys Asn Ile Val
65                  70                  75                  80

aag tat ttg gga tca ttg aag aca aag agc cac ctc cat att att ttg     288
Lys Tyr Leu Gly Ser Leu Lys Thr Lys Ser His Leu His Ile Ile Leu
                85                  90                  95
```

```
gag tat gtg gag aat ggc tca ctt gct aat att atc aag cca aac aaa      336
Glu Tyr Val Glu Asn Gly Ser Leu Ala Asn Ile Ile Lys Pro Asn Lys
            100                 105                 110

ttc gga cct ttt cct gaa tct ttg gtg gct gta tac att gct cag gtg      384
Phe Gly Pro Phe Pro Glu Ser Leu Val Ala Val Tyr Ile Ala Gln Val
        115                 120                 125

ttg gaa ggt ctt gtt tat ctg cat gaa caa ggt gtc att cat aga gat      432
Leu Glu Gly Leu Val Tyr Leu His Glu Gln Gly Val Ile His Arg Asp
    130                 135                 140

atc aag ggc gca aat ata ctg act acc aaa gag ggc ctt gtc aaa ctt      480
Ile Lys Gly Ala Asn Ile Leu Thr Thr Lys Glu Gly Leu Val Lys Leu
145                 150                 155                 160

gct gac ttt gga gtt gcc act aaa ttg act gaa gct gac atc aac act      528
Ala Asp Phe Gly Val Ala Thr Lys Leu Thr Glu Ala Asp Ile Asn Thr
                165                 170                 175

cat tct gtg gtc ggc act cca tac tgg atg gcg cct gag gtc atc gaa      576
His Ser Val Val Gly Thr Pro Tyr Trp Met Ala Pro Glu Val Ile Glu
            180                 185                 190

atg tct ggt gtt tgt gct gcc tct gat atc tgg agt gtg ggt tgc acc      624
Met Ser Gly Val Cys Ala Ala Ser Asp Ile Trp Ser Val Gly Cys Thr
        195                 200                 205

gta att gaa ttg ctg aca tgt gtc ccg cca tat tat gat ctc caa cct      672
Val Ile Glu Leu Leu Thr Cys Val Pro Pro Tyr Tyr Asp Leu Gln Pro
    210                 215                 220

atg cct gcc ctg ttc cgt att gtt cag gat gtg cat cca cca ata cca      720
Met Pro Ala Leu Phe Arg Ile Val Gln Asp Val His Pro Pro Ile Pro
225                 230                 235                 240

gaa gga ctg tca cct gag att act gat ttt ctc cgg caa tgt ttt caa      768
Glu Gly Leu Ser Pro Glu Ile Thr Asp Phe Leu Arg Gln Cys Phe Gln
                245                 250                 255

aag gat gcg atg caa agg cct gat gcg aag aca tta ttg atg cat cca      816
Lys Asp Ala Met Gln Arg Pro Asp Ala Lys Thr Leu Leu Met His Pro
            260                 265                 270

tgg ttg cag aac tca aga cgt gct ttg cct gct tcc ctt cgg caa cct      864
Trp Leu Gln Asn Ser Arg Arg Ala Leu Pro Ala Ser Leu Arg Gln Pro
        275                 280                 285

act cca tta agg aat atc gac ggg gat gat gaa agc tca aga ggg gac      912
Thr Pro Leu Arg Asn Ile Asp Gly Asp Asp Glu Ser Ser Arg Gly Asp
    290                 295                 300

aat aac tcg ggt ttt tgc gac aca cca gga gat acc cga gca act att      960
Asn Asn Ser Gly Phe Cys Asp Thr Pro Gly Asp Thr Arg Ala Thr Ile
305                 310                 315                 320

gct tct aat gtt gac cag gtg aat ggg agg aac gga cca aat atg gac     1008
Ala Ser Asn Val Asp Gln Val Asn Gly Arg Asn Gly Pro Asn Met Asp
                325                 330                 335

tct gct gcg caa agt aaa tct gaa gag ctt cat gat gga aat tta gaa     1056
```

```
Ser Ala Ala Gln Ser Lys Ser Glu Glu Leu His Asp Gly Asn Leu Glu
            340                 345                 350

ctt aca gag ggc att agt tcg aac aat gtg gcg ctt gtg aaa gat aat      1104
Leu Thr Glu Gly Ile Ser Ser Asn Asn Val Ala Leu Val Lys Asp Asn
        355                 360                 365

gtg gtc ctt aat aaa gat ccg aca cta gtt tta cat gag aag tta ccg      1152
Val Val Leu Asn Lys Asp Pro Thr Leu Val Leu His Glu Lys Leu Pro
    370                 375                 380

gtg gaa tct tca ttt gga gat gct gat ttg aat ggc aaa gta atg gca      1200
Val Glu Ser Ser Phe Gly Asp Ala Asp Leu Asn Gly Lys Val Met Ala
385                 390                 395                 400

cat gaa cta ttg caa ggt ggt ctg cca agt aag ata gag ctg gag aat      1248
His Glu Leu Leu Gln Gly Gly Leu Pro Ser Lys Ile Glu Leu Glu Asn
                405                 410                 415

aaa gaa agt tca act gtt gag gat ggt gat gta ttc tca ttt cag gct      1296
Lys Glu Ser Ser Thr Val Glu Asp Gly Asp Val Phe Ser Phe Gln Ala
            420                 425                 430

gga aga cag aat att gac ttc caa atg gtg gtt gaa cct tca gtt gtt      1344
Gly Arg Gln Asn Ile Asp Phe Gln Met Val Val Glu Pro Ser Val Val
        435                 440                 445

gag gga ccg aaa cag ctt agt aga ttt agt gac aaa cct gaa gat gcc      1392
Glu Gly Pro Lys Gln Leu Ser Arg Phe Ser Asp Lys Pro Glu Asp Ala
    450                 455                 460

tct ttg gaa gac tta ttt cca cca atc gac aaa cgg gga gat aat ggg      1440
Ser Leu Glu Asp Leu Phe Pro Pro Ile Asp Lys Arg Gly Asp Asn Gly
465                 470                 475                 480

gct gaa cct tca aca tca act act gtc caa gaa ctt caa tat aat ggt      1488
Ala Glu Pro Ser Thr Ser Thr Thr Val Gln Glu Leu Gln Tyr Asn Gly
                485                 490                 495

gta cac aat gga ttc gtg aag ggg ctc aat gcc agg gtg gag aag caa      1536
Val His Asn Gly Phe Val Lys Gly Leu Asn Ala Arg Val Glu Lys Gln
            500                 505                 510

aaa gaa aat gat agc gag tcc atg aat ggt gga aaa cta att gaa ttt      1584
Lys Glu Asn Asp Ser Glu Ser Met Asn Gly Gly Lys Leu Ile Glu Phe
        515                 520                 525

gcc atg cag cta gaa aat att gat gca tcg ggt ttt ggt gaa cat att      1632
Ala Met Gln Leu Glu Asn Ile Asp Ala Ser Gly Phe Gly Glu His Ile
    530                 535                 540

cct gga gag agt ctt ttc cct ctg cag tct gtg gag tac agc aaa ata      1680
Pro Gly Glu Ser Leu Phe Pro Leu Gln Ser Val Glu Tyr Ser Lys Ile
545                 550                 555                 560

gta gca caa ctg aaa cca ggg gaa agt gaa gat gta ata ctg ttg gcg      1728
Val Ala Gln Leu Lys Pro Gly Glu Ser Glu Asp Val Ile Leu Leu Ala
                565                 570                 575

tgc caa aag ctt cta tca atc ttc agt cac cgg cct gag caa aaa cag      1776
Cys Gln Lys Leu Leu Ser Ile Phe Ser His Arg Pro Glu Gln Lys Gln
            580                 585                 590
```

```
atc tat atg gca cag aat ggt ttc ctc ccg ctg atg gaa ctc ctt gaa     1824
Ile Tyr Met Ala Gln Asn Gly Phe Leu Pro Leu Met Glu Leu Leu Glu
        595                 600                 605

att cct aaa aat cgt att tta tgc tca gtt ctg cag ctt atc aac aac     1872
Ile Pro Lys Asn Arg Ile Leu Cys Ser Val Leu Gln Leu Ile Asn Asn
    610                 615                 620

att gta aaa gat agt aca ggc ttc ctg gaa aat gct tgc ctt gtt ggc     1920
Ile Val Lys Asp Ser Thr Gly Phe Leu Glu Asn Ala Cys Leu Val Gly
625                 630                 635                 640

ctt att cca gtg gtg atg aat ttt gca gaa cca aat cgc cca aag gat     1968
Leu Ile Pro Val Val Met Asn Phe Ala Glu Pro Asn Arg Pro Lys Asp
                645                 650                 655

gtt cgg gtg caa gca tcc ttg ttt ctt cag cag ctt tgt caa gcc agc     2016
Val Arg Val Gln Ala Ser Leu Phe Leu Gln Gln Leu Cys Gln Ala Ser
            660                 665                 670

acc ttg aca ttg caa atg ttc att gca tgt caa ggg atc cct gtt ttg     2064
Thr Leu Thr Leu Gln Met Phe Ile Ala Cys Gln Gly Ile Pro Val Leu
        675                 680                 685

gta agc ttt atg gag cct gac tat gca aaa tac agg gat atg gtt cat     2112
Val Ser Phe Met Glu Pro Asp Tyr Ala Lys Tyr Arg Asp Met Val His
    690                 695                 700

ctt gca att gac ggc att tgg cag gtc ttt aag ctc cag cac tca act     2160
Leu Ala Ile Asp Gly Ile Trp Gln Val Phe Lys Leu Gln His Ser Thr
705                 710                 715                 720

cca aga aat gac ttc tgc cgt ata gca gca aaa aat ggg atg ctt cct     2208
Pro Arg Asn Asp Phe Cys Arg Ile Ala Ala Lys Asn Gly Met Leu Pro
                725                 730                 735

agg ctg gtg aat aca cta tat agc ttg aac gaa gca aca cga ttt gcc     2256
Arg Leu Val Asn Thr Leu Tyr Ser Leu Asn Glu Ala Thr Arg Phe Ala
            740                 745                 750

tca gtc tca ggg tca ggt gct tca gta aca cag aat ggc tca acc ccc     2304
Ser Val Ser Gly Ser Gly Ala Ser Val Thr Gln Asn Gly Ser Thr Pro
        755                 760                 765

cga cga aga tct ggc caa tta gac cct tca gtt ctt gaa agt tgc aag     2352
Arg Arg Arg Ser Gly Gln Leu Asp Pro Ser Val Leu Glu Ser Cys Lys
    770                 775                 780

gcg agg ctg gat cat cat cat tca tct ggt tct ttg caa tct cta caa     2400
Ala Arg Leu Asp His His His Ser Ser Gly Ser Leu Gln Ser Leu Gln
785                 790                 795                 800

gca gat gca gat aag cac cac ata tta atg gac tcc tcg tca tct cct     2448
Ala Asp Ala Asp Lys His His Ile Leu Met Asp Ser Ser Ser Ser Pro
                805                 810                 815

agg ttc agc gac aaa aca agc agt aac ttg gag agg aat gag aac gac     2496
Arg Phe Ser Asp Lys Thr Ser Ser Asn Leu Glu Arg Asn Glu Asn Asp
            820                 825                 830

ttg gtt att cgg cca cag agg ctt agt gtt tct gcc gga agg aca tct     2544
```

```
Leu Val Ile Arg Pro Gln Arg Leu Ser Val Ser Ala Gly Arg Thr Ser
        835                 840                 845

aca gac aga tca ccg aag cac ata gaa cta gtt tca aat ggt cat aat     2592
Thr Asp Arg Ser Pro Lys His Ile Glu Leu Val Ser Asn Gly His Asn
    850                 855                 860

aac agt caa aat gac cag gtc cga cct cta ctg agt tta ttg gag aaa     2640
Asn Ser Gln Asn Asp Gln Val Arg Pro Leu Leu Ser Leu Leu Glu Lys
865                 870                 875                 880

gaa cca cct tct cga cat gta tct gga caa ctt gat tat gtt cgc cat     2688
Glu Pro Pro Ser Arg His Val Ser Gly Gln Leu Asp Tyr Val Arg His
                885                 890                 895

ata tct ggt ttg gaa agg cat gaa acc att ctg cca cta tta cat aca     2736
Ile Ser Gly Leu Glu Arg His Glu Thr Ile Leu Pro Leu Leu His Thr
            900                 905                 910

tcg aca gag agg aaa acg aat ggt gaa ctt gac tta atg atg gat ttt     2784
Ser Thr Glu Arg Lys Thr Asn Gly Glu Leu Asp Leu Met Met Asp Phe
        915                 920                 925

gct gag gtt tct aga cat gga agg gag aat ggc cat ctg gac tct agt     2832
Ala Glu Val Ser Arg His Gly Arg Glu Asn Gly His Leu Asp Ser Ser
    930                 935                 940

ata aaa gct tca aat aga gtt cat agt atg aaa tat gca cct tca gca     2880
Ile Lys Ala Ser Asn Arg Val His Ser Met Lys Tyr Ala Pro Ser Ala
945                 950                 955                 960

agt gcg tct aat gaa ggg gca tca act tct gga gct gca tca caa act     2928
Ser Ala Ser Asn Glu Gly Ala Ser Thr Ser Gly Ala Ala Ser Gln Thr
                965                 970                 975

gca tct ggt gtg tta tca gga tca ggg gtg ctc aat tta aga cct ggg     2976
Ala Ser Gly Val Leu Ser Gly Ser Gly Val Leu Asn Leu Arg Pro Gly
            980                 985                 990

agt aca aca tca tca ggc cca tta gct cag atg ttc tct tcc atg agc     3024
Ser Thr Thr Ser Ser Gly Pro Leu Ala Gln Met Phe Ser Ser Met Ser
        995                 1000                1005

gca gat gtg gca cgt gaa tat ctt gag aaa gta gca gat ctt ctt ttg     3072
Ala Asp Val Ala Arg Glu Tyr Leu Glu Lys Val Ala Asp Leu Leu Leu
    1010                1015                1020

gag ttt gca cag gca gac acc gtg gta aaa tct ctg atg gcc agc caa     3120
Glu Phe Ala Gln Ala Asp Thr Val Val Lys Ser Leu Met Ala Ser Gln
1025                1030                1035                1040

agc ctt ctt aca agg gtt ttc caa atg ttt aac aag ata gag cct cct     3168
Ser Leu Leu Thr Arg Val Phe Gln Met Phe Asn Lys Ile Glu Pro Pro
                1045                1050                1055

att ctt ctt aag att ctt agg tgc att aat cat ttg tcg ggt gat cct     3216
Ile Leu Leu Lys Ile Leu Arg Cys Ile Asn His Leu Ser Gly Asp Pro
            1060                1065                1070

aat tgt ctt gag aca ctt caa cgc aca gat gca atc aaa cat ctg ata     3264
Asn Cys Leu Glu Thr Leu Gln Arg Thr Asp Ala Ile Lys His Leu Ile
        1075                1080                1085
```

```
ccg att ctg gag ctt cgt gat gga cct ctg gtt tat caa ata cat agt     3312
Pro Ile Leu Glu Leu Arg Asp Gly Pro Leu Val Tyr Gln Ile His Ser
    1090                1095                1100

gag gtc ctt aat gct ttg ttc aac ctt tgc aaa att aat aaa aga aga     3360
Glu Val Leu Asn Ala Leu Phe Asn Leu Cys Lys Ile Asn Lys Arg Arg
1105                1110                1115                1120

cag gaa caa gca gcc gaa aat gga atc atc cct cac ttg atg aaa ttt     3408
Gln Glu Gln Ala Ala Glu Asn Gly Ile Ile Pro His Leu Met Lys Phe
                1125                1130                1135

gtc atg tca gac tca cct ttg cgg cag tat gcc ctg cct ttg ctg tgt     3456
Val Met Ser Asp Ser Pro Leu Arg Gln Tyr Ala Leu Pro Leu Leu Cys
            1140                1145                1150

gat atg gct cat gct tct cgc aac tct aga gag cag ttg agg gtt cat     3504
Asp Met Ala His Ala Ser Arg Asn Ser Arg Glu Gln Leu Arg Val His
        1155                1160                1165

gga gga cta gat gtg tac tta aat ctg ctg gag gat gat gca tgg gca     3552
Gly Gly Leu Asp Val Tyr Leu Asn Leu Leu Glu Asp Asp Ala Trp Ala
    1170                1175                1180

tgc acg gca tta gat tct att gct gtt tgt ttg gct cac gat aat gat     3600
Cys Thr Ala Leu Asp Ser Ile Ala Val Cys Leu Ala His Asp Asn Asp
1185                1190                1195                1200

cac aga aaa gtg gaa caa gca ctg ttg aaa aaa gag gcc att cag aaa     3648
His Arg Lys Val Glu Gln Ala Leu Leu Lys Lys Glu Ala Ile Gln Lys
                1205                1210                1215

tta gtg aag ttc ttc caa gat tgc cca gaa cag tat ttt ggc cat ata     3696
Leu Val Lys Phe Phe Gln Asp Cys Pro Glu Gln Tyr Phe Gly His Ile
            1220                1225                1230

ttg gac gct ttc ttg aag ata atc acg aaa tct tct agg cta aat aca     3744
Leu Asp Ala Phe Leu Lys Ile Ile Thr Lys Ser Ser Arg Leu Asn Thr
        1235                1240                1245

gca atg gct acc aat ggt tta aca aca ttg ctc att gca aga ctt gac     3792
Ala Met Ala Thr Asn Gly Leu Thr Thr Leu Leu Ile Ala Arg Leu Asp
    1250                1255                1260

cat cga gaa gct att gcc cgc ttg act ctg ctg aaa ctc ata aag gtt     3840
His Arg Glu Ala Ile Ala Arg Leu Thr Leu Leu Lys Leu Ile Lys Val
1265                1270                1275                1280

gtc tat gag cac cat cct cga cca aaa cag ctc ata gta gag aat gac     3888
Val Tyr Glu His His Pro Arg Pro Lys Gln Leu Ile Val Glu Asn Asp
                1285                1290                1295

ctc cct cag aag cta cag aac ctc att gaa gag cgc agg gat gga caa     3936
Leu Pro Gln Lys Leu Gln Asn Leu Ile Glu Glu Arg Arg Asp Gly Gln
            1300                1305                1310

cgg ggc ggc caa caa gtg ttg gtc aaa caa atg gcc acc tca ttg ctg     3984
Arg Gly Gly Gln Gln Val Leu Val Lys Gln Met Ala Thr Ser Leu Leu
        1315                1320                1325

aaa gcg ctg cac atc aat aca gta ttg tga                             4014
```

```
Lys Ala Leu His Ile Asn Thr Val Leu  *
    1330                1335


      <210> 62
      <211> 1337
      <212> PRT
      <213> Zea mays

      <400> 62
Met Ala Ser Arg Gln His Asn Ala Gln Phe His Lys Ser Lys Thr Leu
 1               5                  10                  15
Asp Asn Lys Tyr Met Leu Gly Asp Glu Ile Gly Lys Gly Ala Tyr Gly
            20                  25                  30
Arg Val Tyr Lys Gly Leu Asp Leu Glu Asn Gly Asp Phe Val Ala Ile
        35                  40                  45
Lys Gln Val Ser Leu Glu Asn Ile Pro Gln Glu Asp Leu Asn Ile Ile
    50                  55                  60
Met Gln Glu Ile Asp Leu Leu Lys Asn Leu Asn His Lys Asn Ile Val
65                  70                  75                  80
Lys Tyr Leu Gly Ser Leu Lys Thr Lys Ser His Leu His Ile Ile Leu
                85                  90                  95
Glu Tyr Val Glu Asn Gly Ser Leu Ala Asn Ile Ile Lys Pro Asn Lys
            100                 105                 110
Phe Gly Pro Phe Pro Glu Ser Leu Val Ala Val Tyr Ile Ala Gln Val
        115                 120                 125
Leu Glu Gly Leu Val Tyr Leu His Glu Gln Gly Val Ile His Arg Asp
    130                 135                 140
Ile Lys Gly Ala Asn Ile Leu Thr Thr Lys Glu Gly Leu Val Lys Leu
145                 150                 155                 160
Ala Asp Phe Gly Val Ala Thr Lys Leu Thr Glu Ala Asp Ile Asn Thr
                165                 170                 175
His Ser Val Val Gly Thr Pro Tyr Trp Met Ala Pro Glu Val Ile Glu
            180                 185                 190
Met Ser Gly Val Cys Ala Ala Ser Asp Ile Trp Ser Val Gly Cys Thr
        195                 200                 205
Val Ile Glu Leu Leu Thr Cys Val Pro Pro Tyr Tyr Asp Leu Gln Pro
    210                 215                 220
Met Pro Ala Leu Phe Arg Ile Val Gln Asp Val His Pro Pro Ile Pro
225                 230                 235                 240
Glu Gly Leu Ser Pro Glu Ile Thr Asp Phe Leu Arg Gln Cys Phe Gln
                245                 250                 255
Lys Asp Ala Met Gln Arg Pro Asp Ala Lys Thr Leu Leu Met His Pro
            260                 265                 270
Trp Leu Gln Asn Ser Arg Arg Ala Leu Pro Ala Ser Leu Arg Gln Pro
        275                 280                 285
Thr Pro Leu Arg Asn Ile Asp Gly Asp Asp Glu Ser Ser Arg Gly Asp
    290                 295                 300
Asn Asn Ser Gly Phe Cys Asp Thr Pro Gly Asp Thr Arg Ala Thr Ile
305                 310                 315                 320
Ala Ser Asn Val Asp Gln Val Asn Gly Arg Asn Gly Pro Asn Met Asp
                325                 330                 335
Ser Ala Ala Gln Ser Lys Ser Glu Glu Leu His Asp Gly Asn Leu Glu
            340                 345                 350
Leu Thr Glu Gly Ile Ser Ser Asn Asn Val Ala Leu Val Lys Asp Asn
        355                 360                 365
Val Val Leu Asn Lys Asp Pro Thr Leu Val Leu His Glu Lys Leu Pro
    370                 375                 380
Val Glu Ser Ser Phe Gly Asp Ala Asp Leu Asn Gly Lys Val Met Ala
385                 390                 395                 400
His Glu Leu Leu Gln Gly Gly Leu Pro Ser Lys Ile Glu Leu Glu Asn
                405                 410                 415
```

Lys Glu Ser Ser Thr Val Glu Asp Gly Asp Val Phe Ser Phe Gln Ala
420 425 430
Gly Arg Gln Asn Ile Asp Phe Gln Met Val Val Glu Pro Ser Val Val
435 440 445
Glu Gly Pro Lys Gln Leu Ser Arg Phe Ser Asp Lys Pro Glu Asp Ala
450 455 460
Ser Leu Glu Asp Leu Phe Pro Pro Ile Asp Lys Arg Gly Asp Asn Gly
465 470 475 480
Ala Glu Pro Ser Thr Ser Thr Thr Val Gln Glu Leu Gln Tyr Asn Gly
485 490 495
Val His Asn Gly Phe Val Lys Gly Leu Asn Ala Arg Val Glu Lys Gln
500 505 510
Lys Glu Asn Asp Ser Glu Ser Met Asn Gly Gly Lys Leu Ile Glu Phe
515 520 525
Ala Met Gln Leu Glu Asn Ile Asp Ala Ser Gly Phe Gly Glu His Ile
530 535 540
Pro Gly Glu Ser Leu Phe Pro Leu Gln Ser Val Glu Tyr Ser Lys Ile
545 550 555 560
Val Ala Gln Leu Lys Pro Gly Glu Ser Glu Asp Val Ile Leu Leu Ala
565 570 575
Cys Gln Lys Leu Leu Ser Ile Phe Ser His Arg Pro Glu Gln Lys Gln
580 585 590
Ile Tyr Met Ala Gln Asn Gly Phe Leu Pro Leu Met Glu Leu Leu Glu
595 600 605
Ile Pro Lys Asn Arg Ile Leu Cys Ser Val Leu Gln Leu Ile Asn Asn
610 615 620
Ile Val Lys Asp Ser Thr Gly Phe Leu Glu Asn Ala Cys Leu Val Gly
625 630 635 640
Leu Ile Pro Val Val Met Asn Phe Ala Glu Pro Asn Arg Pro Lys Asp
645 650 655
Val Arg Val Gln Ala Ser Leu Phe Leu Gln Gln Leu Cys Gln Ala Ser
660 665 670
Thr Leu Thr Leu Gln Met Phe Ile Ala Cys Gln Gly Ile Pro Val Leu
675 680 685
Val Ser Phe Met Glu Pro Asp Tyr Ala Lys Tyr Arg Asp Met Val His
690 695 700
Leu Ala Ile Asp Gly Ile Trp Gln Val Phe Lys Leu Gln His Ser Thr
705 710 715 720
Pro Arg Asn Asp Phe Cys Arg Ile Ala Ala Lys Asn Gly Met Leu Pro
725 730 735
Arg Leu Val Asn Thr Leu Tyr Ser Leu Asn Glu Ala Thr Arg Phe Ala
740 745 750
Ser Val Ser Gly Ser Gly Ala Ser Val Thr Gln Asn Gly Ser Thr Pro
755 760 765
Arg Arg Arg Ser Gly Gln Leu Asp Pro Ser Val Leu Glu Ser Cys Lys
770 775 780
Ala Arg Leu Asp His His His Ser Ser Gly Ser Leu Gln Ser Leu Gln
785 790 795 800
Ala Asp Ala Asp Lys His His Ile Leu Met Asp Ser Ser Ser Ser Pro
805 810 815
Arg Phe Ser Asp Lys Thr Ser Ser Asn Leu Glu Arg Asn Glu Asn Asp
820 825 830
Leu Val Ile Arg Pro Gln Arg Leu Ser Val Ser Ala Gly Arg Thr Ser
835 840 845
Thr Asp Arg Ser Pro Lys His Ile Glu Leu Val Ser Asn Gly His Asn
850 855 860
Asn Ser Gln Asn Asp Gln Val Arg Pro Leu Leu Ser Leu Leu Glu Lys
865 870 875 880
Glu Pro Pro Ser Arg His Val Ser Gly Gln Leu Asp Tyr Val Arg His
885 890 895
Ile Ser Gly Leu Glu Arg His Glu Thr Ile Leu Pro Leu Leu His Thr
900 905 910

Ser Thr Glu Arg Lys Thr Asn Gly Glu Leu Asp Leu Met Met Asp Phe
915 920 925
Ala Glu Val Ser Arg His Gly Arg Glu Asn Gly His Leu Asp Ser Ser
930 935 940
Ile Lys Ala Ser Asn Arg Val His Ser Met Lys Tyr Ala Pro Ser Ala
945 950 955 960
Ser Ala Ser Asn Glu Gly Ala Ser Thr Ser Gly Ala Ala Ser Gln Thr
965 970 975
Ala Ser Gly Val Leu Ser Gly Ser Gly Val Leu Asn Leu Arg Pro Gly
980 985 990
Ser Thr Thr Ser Ser Gly Pro Leu Ala Gln Met Phe Ser Ser Met Ser
995 1000 1005
Ala Asp Val Ala Arg Glu Tyr Leu Glu Lys Val Ala Asp Leu Leu Leu
1010 1015 1020
Glu Phe Ala Gln Ala Asp Thr Val Val Lys Ser Leu Met Ala Ser Gln
1025 1030 1035 1040
Ser Leu Leu Thr Arg Val Phe Gln Met Phe Asn Lys Ile Glu Pro Pro
1045 1050 1055
Ile Leu Leu Lys Ile Leu Arg Cys Ile Asn His Leu Ser Gly Asp Pro
1060 1065 1070
Asn Cys Leu Glu Thr Leu Gln Arg Thr Asp Ala Ile Lys His Leu Ile
1075 1080 1085
Pro Ile Leu Glu Leu Arg Asp Gly Pro Leu Val Tyr Gln Ile His Ser
1090 1095 1100
Glu Val Leu Asn Ala Leu Phe Asn Leu Cys Lys Ile Asn Lys Arg Arg
1105 1110 1115 1120
Gln Glu Gln Ala Ala Glu Asn Gly Ile Ile Pro His Leu Met Lys Phe
1125 1130 1135
Val Met Ser Asp Ser Pro Leu Arg Gln Tyr Ala Leu Pro Leu Leu Cys
1140 1145 1150
Asp Met Ala His Ala Ser Arg Asn Ser Arg Glu Gln Leu Arg Val His
1155 1160 1165
Gly Gly Leu Asp Val Tyr Leu Asn Leu Leu Glu Asp Asp Ala Trp Ala
1170 1175 1180
Cys Thr Ala Leu Asp Ser Ile Ala Val Cys Leu Ala His Asp Asn Asp
1185 1190 1195 1200
His Arg Lys Val Glu Gln Ala Leu Leu Lys Lys Glu Ala Ile Gln Lys
1205 1210 1215
Leu Val Lys Phe Phe Gln Asp Cys Pro Glu Gln Tyr Phe Gly His Ile
1220 1225 1230
Leu Asp Ala Phe Leu Lys Ile Ile Thr Lys Ser Ser Arg Leu Asn Thr
1235 1240 1245
Ala Met Ala Thr Asn Gly Leu Thr Thr Leu Leu Ile Ala Arg Leu Asp
1250 1255 1260
His Arg Glu Ala Ile Ala Arg Leu Thr Leu Leu Lys Leu Ile Lys Val
1265 1270 1275 1280
Val Tyr Glu His His Pro Arg Pro Lys Gln Leu Ile Val Glu Asn Asp
1285 1290 1295
Leu Pro Gln Lys Leu Gln Asn Leu Ile Glu Glu Arg Arg Asp Gly Gln
1300 1305 1310
Arg Gly Gly Gln Gln Val Leu Val Lys Gln Met Ala Thr Ser Leu Leu
1315 1320 1325
Lys Ala Leu His Ile Asn Thr Val Leu
1330 1335

<210> 63
<211> 20
<212> DNA
<213> Zea mays

<400> 63
atggcgtcga ggcagcacaa 20

<210> 64
<211> 20
<212> DNA
<213> Zea mays

<400> 64
tcacaatact gtattgatgt 20

<210> 65
<211> 1196
<212> DNA
<213> Zea mays

<220>
<221> CDS
<222> (117)...(924)

<400> 65
ccaaaatcga ccttgtcgac ccggatccga gctaggcacc cgcacgacca cccgcctcct 60
gccggtcacc gggacctagc atcgccgggc ggcggcgggg gaggggaagg taggcg atg 119
Met
1

agg aaa cgt gag cgg gag aac ccg tgc ggg atc tgc ggg cac tac cac 167
Arg Lys Arg Glu Arg Glu Asn Pro Cys Gly Ile Cys Gly His Tyr His
5 10 15

aag tac gag gag ggg gag gta tgc ggg gtg tgc ggc cac cgc tgg aaa 215
Lys Tyr Glu Glu Gly Glu Val Cys Gly Val Cys Gly His Arg Trp Lys
20 25 30

ccg tcg gac ggg gag ggc gcc ccg tcg agg cac gag tcc gcc ttc ccc 263
Pro Ser Asp Gly Glu Gly Ala Pro Ser Arg His Glu Ser Ala Phe Pro
35 40 45

aca gag gtg ctc aag gac ttt ctc ttc ctc ggg agc tac gac aac gcc 311
Thr Glu Val Leu Lys Asp Phe Leu Phe Leu Gly Ser Tyr Asp Asn Ala
50 55 60 65

tca cgc tcc gag gtc ctc aga acg ctc agt atc tcc cac atc ctt aat 359
Ser Arg Ser Glu Val Leu Arg Thr Leu Ser Ile Ser His Ile Leu Asn
70 75 80

act gtg cct gat tgc cac aat ctg tac aag aat tca ttc aca tat cac 407
Thr Val Pro Asp Cys His Asn Leu Tyr Lys Asn Ser Phe Thr Tyr His
85 90 95

agt ctt caa cgg gac agg cct ttg gat ttt gat gag gca aac caa tat 455
Ser Leu Gln Arg Asp Arg Pro Leu Asp Phe Asp Glu Ala Asn Gln Tyr
100 105 110

cta gag aaa tgt gaa aga gat aaa tcg cgt gtt ctt gtt cat tgc atg 503
Leu Glu Lys Cys Glu Arg Asp Lys Ser Arg Val Leu Val His Cys Met
115 120 125

act gga aaa aac agg tct gct gct att gtt gca gcc ttc ttg atg aag 551
Thr Gly Lys Asn Arg Ser Ala Ala Ile Val Ala Ala Phe Leu Met Lys
130 135 140 145

tct aga ggg tgg aga ctt gca caa tct ttt caa tgg gtg aaa gat cga 599
Ser Arg Gly Trp Arg Leu Ala Gln Ser Phe Gln Trp Val Lys Asp Arg

```
                150                 155                 160

aga cca cag gtt caa ctg aca gat gct tct cag aag gag ctt ttg gag      647
Arg Pro Gln Val Gln Leu Thr Asp Ala Ser Gln Lys Glu Leu Leu Glu
            165                 170                 175

tat gag ctg aag ctt ttc gga ccc agt act caa acc ttg gtt ccc act      695
Tyr Glu Leu Lys Leu Phe Gly Pro Ser Thr Gln Thr Leu Val Pro Thr
        180                 185                 190

gaa tca ttt gct tct ctt gga ttt ggt tat aca aaa cca gca ggt gac      743
Glu Ser Phe Ala Ser Leu Gly Phe Gly Tyr Thr Lys Pro Ala Gly Asp
    195                 200                 205

atg caa acg cct gtg ttc aac caa acg acc atg cca tcc ata ttc gag      791
Met Gln Thr Pro Val Phe Asn Gln Thr Thr Met Pro Ser Ile Phe Glu
210                 215                 220                 225

cga ttc aac cca aat gat cgc ccc acc aat ttc act ttt gga gtt gag      839
Arg Phe Asn Pro Asn Asp Arg Pro Thr Asn Phe Thr Phe Gly Val Glu
                230                 235                 240

agg aaa gca gaa gct aac cct cga gat aac aat agt gat ggc agg gta      887
Arg Lys Ala Glu Ala Asn Pro Arg Asp Asn Asn Ser Asp Gly Arg Val
            245                 250                 255

aac cca act tca acg gag aac ctg atg gac agc tct t aagatgattt         934
Asn Pro Thr Ser Thr Glu Asn Leu Met Asp Ser Ser
        260                 265

cacgatgtcg ctgttgaatc tccttgttcc agaagtaaac ccgctgtaat ctgcatgtcg    994
cggaagtaaa cctactgctc ggtgcctcgg tgcaccgaat acttatccgg tgtctaggtt   1054
tcagggaggt tttgcccttg tcgcatgaga aatggttctc tggaaaccga tctttgcttt   1114
agaatgactg tatgaaatag atgtttgtgg ctaaaaaaaa aaaaaaaaaa aaaaaaaaaa   1174
aaaaaaaaaa aaaaaaaaaa aa                                            1196

      <210> 66
      <211> 269
      <212> PRT
      <213> Zea mays

      <400> 66
Met Arg Lys Arg Glu Arg Glu Asn Pro Cys Gly Ile Cys Gly His Tyr
 1               5                  10                  15
His Lys Tyr Glu Glu Gly Glu Val Cys Gly Val Cys Gly His Arg Trp
            20                  25                  30
Lys Pro Ser Asp Gly Glu Gly Ala Pro Ser Arg His Glu Ser Ala Phe
        35                  40                  45
Pro Thr Glu Val Leu Lys Asp Phe Leu Phe Leu Gly Ser Tyr Asp Asn
    50                  55                  60
Ala Ser Arg Ser Glu Val Leu Arg Thr Leu Ser Ile Ser His Ile Leu
65                  70                  75                  80
Asn Thr Val Pro Asp Cys His Asn Leu Tyr Lys Asn Ser Phe Thr Tyr
                85                  90                  95
His Ser Leu Gln Arg Asp Arg Pro Leu Asp Phe Asp Glu Ala Asn Gln
            100                 105                 110
Tyr Leu Glu Lys Cys Glu Arg Asp Lys Ser Arg Val Leu Val His Cys
        115                 120                 125
Met Thr Gly Lys Asn Arg Ser Ala Ala Ile Val Ala Ala Phe Leu Met
    130                 135                 140
Lys Ser Arg Gly Trp Arg Leu Ala Gln Ser Phe Gln Trp Val Lys Asp
145                 150                 155                 160
```

```
Arg Arg Pro Gln Val Gln Leu Thr Asp Ala Ser Gln Lys Glu Leu Leu
                165                 170                 175
Glu Tyr Glu Leu Lys Leu Phe Gly Pro Ser Thr Gln Thr Leu Val Pro
            180                 185                 190
Thr Glu Ser Phe Ala Ser Leu Gly Phe Gly Tyr Thr Lys Pro Ala Gly
        195                 200                 205
Asp Met Gln Thr Pro Val Phe Asn Gln Thr Thr Met Pro Ser Ile Phe
    210                 215                 220
Glu Arg Phe Asn Pro Asn Asp Arg Pro Thr Asn Phe Thr Phe Gly Val
225                 230                 235                 240
Glu Arg Lys Ala Glu Ala Asn Pro Arg Asp Asn Asn Ser Asp Gly Arg
                245                 250                 255
Val Asn Pro Thr Ser Thr Glu Asn Leu Met Asp Ser Ser
            260                 265
```

<210> 67
<211> 25
<212> DNA
<213> Zea mays

<400> 67
atgaggaaac gtgagcggga gaacc 25

<210> 68
<211> 25
<212> DNA
<213> Zea mays

<400> 68
ttaagagctg tccatcaggt tctcc 25

<210> 69
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Designed oligonucleotide based upon the adapter sequence and poly T to remove clones which have a poly A tail but no cDNA.

<400> 69
tcgacccacg cgtccgaaaa aaaaaaaaaa aaaaaa 36

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
23 November 2000 (23.11.2000)

PCT

(10) International Publication Number
WO 00/70059 A3

(51) International Patent Classification[7]: C12N 15/54, 15/29, 15/82, 9/12, A01H 5/00

(21) International Application Number: PCT/US00/11687

(22) International Filing Date: 28 April 2000 (28.04.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/134,292 14 May 1999 (14.05.1999) US
60/142,996 8 July 1999 (08.07.1999) US

(71) Applicant *(for all designated States except US)*: **PIONEER HI-BRED INTERNATIONAL, INC.** [US/US]; 800 Capital Square, 400 Locust Street, Des Moines, IA 50309 (US).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: **HELENTJARIS, Timothy, G.** [US/US]; 2960 NW 73rd Lane, Ankeny, IA 50021 (US).

(74) Agents: **VARLEY, Karen, K.** et al.; 7100 N.W. 62nd Avenue, Darwin Building, Johnston, IA 50131-1000 (US).

(81) Designated States *(national)*: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *With international search report.*

(88) Date of publication of the international search report:
5 April 2001

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 00/70059 A3

(54) Title: SIGNAL TRANSDUCTION GENES AND METHODS OF USE

(57) Abstract: The invention provides isolated signal transduction nucleic acids and their encoded proteins. The present invention provides methods and compositions relating to altering signal transduction gene expression levels in plants. The invention further provides recombinant expression cassettes, host cells, transgenic plants, and antibody compositions.

INTERNATIONAL SEARCH REPORT

Interr al Application No
PCT/US 00/11687

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/54 C12N15/29 C12N15/82 C12N9/12 A01H5/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N A01H

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
WPI Data, BIOSIS, EPO-Internal, EMBL, PAJ

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE EMBL [Online] ACCESSION NO: D26601, 23 February 1994 (1994-02-23) MACHIDA, Y., ET AL.: "Tobacco mRNA for protein kinase, complete cds." XP002147429 the whole document	1,13
X	-& BANNO, H., ET AL.: "NPK1, a tobacco gene that encodes a protein with a domain homologous to yeast BCK1, STE11, and Byr 2 protein kinases." MOLECULAR AND CELLULAR BIOLOGY, vol. 13, 1993, pages 4745-4752, XP000946162 the whole document	1-3,13
	--- -/--	

[X] Further documents are listed in the continuation of box C. [X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
18 December 2000	22.12.00

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Maddox, A

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE WPI Section Ch, Week 199336 Derwent Publications Ltd., London, GB; Class C06, AN 1993-284683 XP002147431 & JP 05 199885 A (SUMITOMO CHEM CO LTD), 10 August 1993 (1993-08-10) abstract -& DATABASE EMBL [Online] ACCESSION NO: E05289, 8 October 1997 (1997-10-08) "DNA encoding a protein kinase that is homologous to yeast STE11." XP002147430 the whole document	1-13

X	DATABASE EMBL [Online] ACCESSION NO:AI600545, 26 April 1999 (1999-04-26) WALBOT, V.: "486076F02.x1 486 - leaf primordia cDNA library from Hake lab Zea mays cDNA,mRNA sequence." XP002147371 see sequence	1

X	DATABASE DBEST [Online] dbEST Id:1200941, 11 August 1997 (1997-08-11) SASAKI, T.: XP002147372	1,13
Y	see whole document & DATABASE EMBL [Online] ACCESSION NO:C22570, 15 August 1997 (1997-08-15) SASAKI, T., ET AL.: "Rice cDNA, partial sequence (E40447_2A)." abstract	2-12

X	DATABASE EMBL [Online] ACCESSION NO:Y14316, 4 August 1997 (1997-08-04) HENRY Y.: "Arabidopsis thaliana mRNA for MAP3K gamma protein kinase" XP002147373	1,13
Y	the whole document -& JOUANNIC, S., ET AL.: "Characterisation of novel plant genes encoding MEKK/STE11 and RAF-related protein kinases." GENE, vol. 229, 18 March 1999 (1999-03-18), pages 171-181, XP004161172 the whole document	2-12

	-/--	

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 98 18939 A (UNIV RUTGERS) 7 May 1998 (1998-05-07) page 32 -page 36	2-12
X	DATABASE EMBL [Online] ACCESSION NO: AB007790, 14 October 1998 (1998-10-14) YAMAGUCHI-SHINOZAKI K.;: "Arabidopsis thaliana mRNA for DREB2A, complete cds." XP002155762 the whole document	1,13
X	DATABASE EMBL [Online] ACCESSION NO: AB010692, 3 February 1998 (1998-02-03) NAKAMURA, Y-. ET AL.: "Arabidopsis thaliana genomic DNA, chromosome 5, TAC clone:K18I23." XP002155763 nts 53956-54963	1-3,13
X	DATABASE EMBL [Online] ACCESSION NO: D23249, 28 November 1993 (1993-11-28) SASAKI, T., ET AL.: " Rice cDNA, partial sequence (C2491_1A)." XP002155764 the whole document	1,13
P,X	DATABASE EMBL [Online] ACCESSION NO: AI668456, 17 May 1999 (1999-05-17) WALBOT, V., ET AL.: "mRNA sequence." XP002155765 the whole document	1,13
P,X	DATABASE EMBL [Online] ACCESSION NO: AW395576, 8 February 2000 (2000-02-08) SHOEMAKER, R., ET AL.: "sg72g10.y1 Gm-c1007 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:Gm-c1007-2011 5' similar to TR:O82132 O82132 DREB2A. ;, mRNA sequence." XP002155766 the whole document	1,13
P,X	DATABASE EMBL [Online] ACCESSION NO:AW076331, 18 October 1999 (1999-10-18) WALBOT, V.: "614065F09.y1 614 - root cDNA library from Walbot Lab Zea mays cDNA, mRNA sequence." XP002147374 the whole document	1,13

-/--

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	DATABASE WPI Section Ch, Week 200034 Derwent Publications Ltd., London, GB; Class C06, AN 2000-389819 XP002155767 & JP 2000 116260 A (BIO-ORIENTED TECHNOLOGY RES ADVANCEMENT), 25 April 2000 (2000-04-25) abstract & DATABASE GENESEQ [Online] ACCESSION NO: A53412, 4 October 2000 (2000-10-04) abstract	1-3,13
P,X	MENKE FRANK L H ET AL: "A novel jasmonate- and elicitor-responsive element in the periwinkle secondary metabolite biosynthetic gene Str interacts with a jasmonate- and elicitor-inducible AP2-domain transcription factor, ORCA2." EMBO (EUROPEAN MOLECULAR BIOLOGY ORGANIZATION) JOURNAL, vol. 18, no. 16, 16 August 1999 (1999-08-16), pages 4455-4463, XP002155761 ISSN: 0261-4189 figure 4	13
P,X	WO 00 08187 A (VERBRUGGEN NATHALIE ;VLAAMS INTERUNIV INST BIOTECH (BE); LEE JEONG) 17 February 2000 (2000-02-17) see SEQ ID NO:33 page 143 -page 146	1-13
P,X	WO 00 09724 A (GEN HOSPITAL CORP) 24 February 2000 (2000-02-24) see figs 14 and 15 the whole document	1-13
E	WO 00 46383 A (UNIV LEIDEN ;MEMELINK JOHAN (NL); FITS CORNELIA THEODORA ELISABE () 10 August 2000 (2000-08-10) see SEQ ID NO:1	1,13
E	EP 1 033 405 A (CERES INC) 6 September 2000 (2000-09-06) see SEQ ID NOS:3964 and 3965	1,13
A	DATABASE EMBL [Online] ACCESSION NO:AB016801, 9 February 1999 (1999-02-09) KUSANO T.: "Zea mays mRNA for MAP kinase 4, complete cds." XP002147375 the whole document	1-13

-/--

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	DATABASE EMBL [Online] ACCESSION NO:AB016802, 9 February 1999 (1999-02-09) KUSANO T.: "Zea mays mRNA for MAP kinase 5, complete cds." XP002147376 the whole document ---	1-13
A	HARDIN SHANE C ET AL: "Molecular cloning and characterization of maize ZmMEK1, a protein kinase with a catalytic domain homologous to mitogen- and stress-activated protein kinase kinases." PLANTA (BERLIN), vol. 206, no. 4, November 1998 (1998-11), pages 577-584, XP000939378 ISSN: 0032-0935 the whole document & DATABASE EMBL [Online] ACCESSION NO:U83625, 8 January 1998 (1998-01-08) ---	1-13
A	DATABASE WPI Section Ch, Week 199720 Derwent Publications Ltd., London, GB; Class B04, AN 1997-220416 XP002079552 & JP 09 065881 A (NORINSUISANSHO NOGYO SEIBUTSU SHIGEN), 11 March 1997 (1997-03-11) abstract -& DATABASE GENESEQ [Online] ACCESSION NO:T60349, 4 June 1997 (1997-06-04) "MAP kinase #1 coding sequence" XP002079541 the whole document -& DATABASE GENESEQ [Online] ACCESSION NO:W15512, 4 June 1997 (1997-06-04) "MAP kinase #1" XP002079540 the whole document ---	1-13
A	SEO ET AL: "Tobacco MAP kinase: a possible mediator in wound signal transduction pathways" SCIENCE,AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE,,US, vol. 270, 22 December 1995 (1995-12-22), pages 1988-1992, XP002079542 ISSN: 0036-8075 the whole document --- -/--	1-13

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	LIGTERINK ET AL: "Receptor-mediated activation of a MAP kinase in pathogen defense of plants" SCIENCE,US,AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE,, vol. 276, 27 June 1997 (1997-06-27), pages 2054-2057, XP002079548 ISSN: 0036-8075 the whole document	1-13

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 00/11687

Box I Observations where certain claims wer found uns archable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of c rtain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☒ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

1-13 all partially, subject groups 1 and 3.

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

R mark on Pr test

☐ The additional search fees wer accompanied by th applicant's protest.

☒ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1-13 all partially

Maize nucleic acid and corresponding protein as characterised by SEQ ID NOS:1 and 2, polynucleotides and proteins related to said sequences as defined by the claims, recombinant expression cassettes,host cells,transgenic plants,methods for modulation of signal transduction gene expression, based on said SEQ ID NOS:1 and 2 and said related polynucleotides and proteins.

2. Claims: 1-13 all partially

Maize nucleic acid and corresponding protein as characterised by SEQ ID NOS:5 and 6, polynucleotides and proteins related to said sequences as defined by the claims, recombinant expression cassettes,host cells,transgenic plants,methods for modulation of signal transduction gene expression, based on said SEQ ID NOS:5 and 6 and said related polynucleotides and proteins

3. Claims: 1-13 all partially

Maize nucleic acid and corresponding protein as characterised by SEQ ID NOS:9 and 10, polynucleotides and proteins related to said sequences as defined by the claims, recombinant expression cassettes,host cells,transgenic plants,methods for modulation of signal transduction gene expression, based on said SEQ ID NOS:9 and 10 and said related polynucleotides and proteins.

4. Claims: 1-13 all partially

Maize nucleic acid and corresponding protein as characterised by SEQ ID NOS:13 and 14, polynucleotides and proteins related to said sequences as defined by the claims, recombinant expression cassettes,host cells,transgenic plants,methods for modulation of signal transduction gene expression, based on said SEQ ID NOS:13 and 14 and said related polynucleotides and proteins.

5. Claims: 1-13 all partially

Maize nucleic acid and corresponding protein as characterised by SEQ ID NOS:17 and 18, polynucleotides and proteins related to said sequences as defined by the claims, recombinant expression cassettes,host cells,transgenic plants,methods for modulation of signal transduction gene expression, based on said SEQ ID NOS:17 and 18 and said

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

related polynucleotides and proteins.

6. Claims: 1-13 all partially

Maize nucleic acid and corresponding protein as characterised by SEQ ID NOS:21 and 22, polynucleotides and proteins related to said sequences as defined by the claims, recombinant expression cassettes,host cells,transgenic plants,methods for modulation of signal transduction gene expression, based on said SEQ ID NOS:21 and 22 and said related polynucleotides and proteins.

7. Claims: 1-13 all partially

Maize nucleic acid and corresponding protein as characterised by SEQ ID NOS:25 and 26, polynucleotides and proteins related to said sequences as defined by the claims, recombinant expression cassettes,host cells,transgenic plants,methods for modulation of signal transduction gene expression, based on said SEQ ID NOS:25 and 26 and said related polynucleotides and proteins.

8. Claims: 1-13 all partially

Maize nucleic acid and corresponding protein as characterised by SEQ ID NOS:29 and 30, polynucleotides and proteins related to said sequences as defined by the claims, recombinant expression cassettes,host cells,transgenic plants,methods for modulation of signal transduction gene expression, based on said SEQ ID NOS:29 and 30 and said related polynucleotides and proteins.

9. Claims: 1-13 all partially

Maize nucleic acid and corresponding protein as characterised by SEQ ID NOS:33 and 34, polynucleotides and proteins related to said sequences as defined by the claims, recombinant expression cassettes,host cells,transgenic plants,methods for modulation of signal transduction gene expression, based on said SEQ ID NOS:33 and 34 and said related polynucleotides and proteins.

10. Claims: 1-13 all partially

Maize nucleic acid and corresponding protein as characterised by SEQ ID NOS:37 and 38, polynucleotides and proteins related to said sequences as defined by the claims, recombinant expression cassettes,host cells,transgenic plants,methods for modulation of signal transduction gene expression, based on said SEQ ID NOS:37 and 38 and said

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

related polynucleotides and proteins.

11. Claims: 1-13 all partially

Maize nucleic acid and corresponding protein as characterised by SEQ ID NOS:41 and 42, polynucleotides and proteins related to said sequences as defined by the claims, recombinant expression cassettes,host cells,transgenic plants,methods for modulation of signal transduction gene expression, based on said SEQ ID NOS:41 and 42 and said related polynucleotides and proteins.

12. Claims: 1-13 all partially

Maize nucleic acid and corresponding protein as characterised by SEQ ID NOS:45 and 46, polynucleotides and proteins related to said sequences as defined by the claims, recombinant expression cassettes,host cells,transgenic plants,methods for modulation of signal transduction gene expression, based on said SEQ ID NOS:45 and 46 and said related polynucleotides and proteins.

13. Claims: 1-13 all partially

Maize nucleic acid and corresponding protein as characterised by SEQ ID NOS:49 and 50, polynucleotides and proteins related to said sequences as defined by the claims, recombinant expression cassettes,host cells,transgenic plants,methods for modulation of signal transduction gene expression, based on said SEQ ID NOS:49 and 50 and said related polynucleotides and proteins.

14. Claims: 1-13 all partially

Maize nucleic acid and corresponding protein as characterised by SEQ ID NOS:53 and 54, polynucleotides and proteins related to said sequences as defined by the claims, recombinant expression cassettes,host cells,transgenic plants,methods for modulation of signal transduction gene expression, based on said SEQ ID NOS:53 and 54 and said related polynucleotides and proteins.

15. Claims: 1-13 all partially

Maize nucleic acid and corresponding protein as characterised by SEQ ID NOS:57 and 58, polynucleotides and proteins related to said sequences as defined by the claims, recombinant expression cassettes,host cells,transgenic plants,methods for modulation of signal transduction gene expression, based on said SEQ ID NOS:57 and 58 and said

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

related polynucleotides and proteins.

16. Claims: 1-13 all partially

Maize nucleic acid and corresponding protein as characterised by SEQ ID NOS:61 and 62, polynucleotides and proteins related to said sequences as defined by the claims, recombinant expression cassettes,host cells,transgenic plants,methods for modulation of signal transduction gene expression, based on said SEQ ID NOS:61 and 62 and said related polynucleotides and proteins.

17. Claims: 1-13 all partially

Maize nucleic acid and corresponding protein as characterised by SEQ ID NOS:65 and 66, polynucleotides and proteins related to said sequences as defined by the claims, recombinant expression cassettes,host cells,transgenic plants,methods for modulation of signal transduction gene expression, based on said SEQ ID NOS:65 and 66 and said related polynucleotides and proteins.

INTERNATIONAL SEARCH REPORT

Information on patent family members

Inter al Application No
PCT/US 00/11687

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
JP 5199885 A	10-08-1993	NONE	
WO 9818939 A	07-05-1998	US 5977442 A	02-11-1999
		AU 4916897 A	22-05-1998
JP 2000116260 A	25-04-2000	NONE	
WO 0008187 A	17-02-2000	AU 5419799 A	28-02-2000
WO 0009724 A	24-02-2000	AU 5475399 A	06-03-2000
		AU 5553599 A	06-03-2000
		WO 0009656 A	24-02-2000
WO 0046383 A	10-08-2000	AU 2580600 A	25-08-2000
EP 1033405 A	06-09-2000	NONE	
JP 9065881 A	11-03-1997	JP 2945953 B	06-09-1999